The Dissertation Committee for Daniel Champlin Propheter Certifies that this is the approved version of the following dissertation:

Advances in Protein Microarray Technology for Glycomic Analysis

Committee:

Brent L. Iverson, Co-Supervisor

Lara K. Mahal, Co-Supervisor

Dionicio R. Siegel

Adrian T. Keatinge-Clay

Walter L. Fast

Advances in Protein Microarray Technology for Glycomic Analysis

by

Daniel Champlin Propheter, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2011

Dedication

This dissertation is dedicated to my mother Laura, and my two older brothers Stephen and Geoff, and sister-in-law Romina, I am grateful for all of your sacrifices. I would also like to dedicate this to my extended family, Paula, Mark, and Shawn, and to those family members that are no longer with us: Opa, Nanita, Nana, Grandma, and Herb. I thank you all for your love, support, and belief in me.

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Lara K. Mahal for her guidance, and more importantly, patience throughout my graduate career. She has been an excellent mentor and has taught me many techniques and skills that have made me the scientist I am today. I want to thank the members of my committee for their comments and insightful questions into the research presented herein. I would also like to thank members of the Mahal lab for their continuing support, help, and patience throughout my time in the group. Specifically, I would like to thank former members Dr. Ku-Lung Hsu and Dr. Lakshmi Krishnamoorthy, and current member Dr. John F. Rakus for their advice, criticisms, and suggestions at both UT Austin and NYU. I would also like to thank my undergraduate adviser at Regis University, Dr. Denise E. Guinn, for her guidance when I first started working independently in a chemistry lab. I would also like to thank my colleagues and friends I have made at UT Austin, NYU, and Albert Einstein School of Medicine for many helpful scientific discussions. I would also like to thank my friend Nancy Hom, the only person and/or object at NYU that hasn't given me an ulcer. For all those I cannot recall at the moment, and for those mentioned above, I thank you for your generosity and support through the years.

Advances in Protein Microarray Technology for Glycomic Analysis

Daniel Champlin Propheter, Ph.D.

The University of Texas at Austin, 2011

Co-Supervisor: Brent L. Iverson

Co-Supervisor: Lara K. Mahal

The cell surface is enveloped with a myriad of carbohydrates that form complex matrices of oligosaccharides. Carbohydrate recognition plays crucial and varying roles in cellular trafficking, differentiation, and bacterial pathogenesis. Lectin microarray technology presents a unique platform for the high-throughput analysis of these structurally diverse classes of biopolymers. One significant hinderance of this technology has been the limitation imposed by the set of commercially available plant lectins used in the array. To enhance the reproducibility and scope of the lectin panel, our lab generated a small set of bacteria-derived recombinant lectins.

This dissertation describes the unique advantages that recombinant lectins have over traditional plant-derived lectins. The recombinant lectins are expressed with a common fusion tag, glutathione-*S*-transferase (GST), which can be used as an immobilization handle on glutathione (GSH)-modified substrates. Although protein immobilization *via* fusion tags in a microarray format is not novel, our work demonstrates that protein activity through site-specific immobilization is enhanced when the protein is properly oriented. Although orientation enhanced the activity of our GST-

tagged recombinant lectins, the GSH-surface modification precluded the printing of non-GST-tagged lectins, such as the traditional plant lectins, thus limiting the structural resolution of our arrays. To solve this issue, we developed a novel print technique which allows the one-step deposition and orientation of GST-tagged proteins in a microarray format. To expand our view of the glycome, we further adapt this method for the *in situ* orientation of unmodified IgG and IgM antibodies using GST-tagged antibody-binding proteins.

Another advantage of recombinant lectins is in the ease of genomic manipulation, wherein we could tailor the binding domain to bind a different antigen. We demonstrate this by producing non-binding variants of the recombinant lectins to act as negative controls in our microarrays. Along with the non-binding variants, we developed a lectin displayed on the surface of phage. In the hopes generating more novel lectins, I will describe our current efforts of lectin evolution using phage-displayed GafD. By generating novel tools in lectin microarray technology, we enhance our understanding of the role of carbohydrates on a global scale.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to current microarray technology

1.1 Introduction

Carbohydrates form complex, cellular matrices on the surfaces of a wide-range of species including mammals and pathogenic bacteria (*1 - 5*). Carbohydrates mediate key intracellular interactions (e.g., protein trafficking) and intercellular interactions (e.g., host-pathogen symbiosis) (Figure 1.1). It is currently estimated that ~50% of all mammalian proteins are glycosylated (*6*). Abnormal changes in the glycosylation machinery can often lead to severe physical defects, classes of which are genetically-linked, so called congenital disorders which affect the *N*-linked glycosylation pathway (*3, 7*).

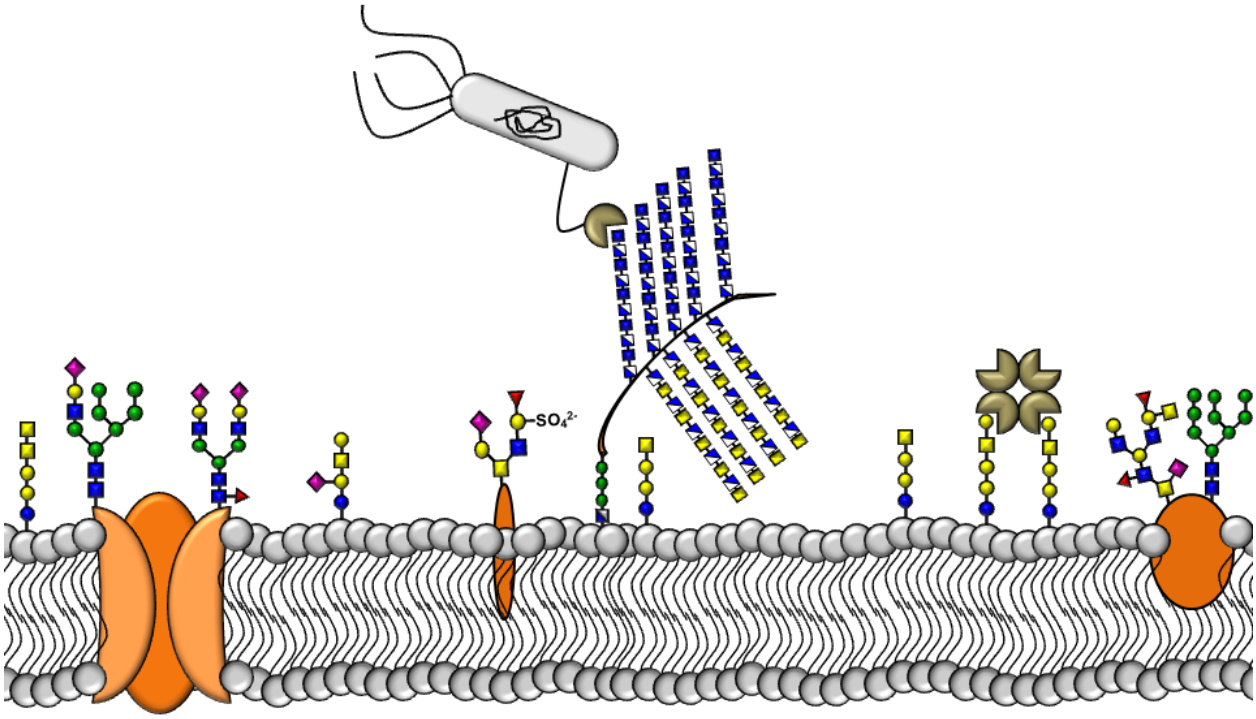

Figure 1.1 Cell-surface glycans as important mediators in cellular functions. Carbohydrates are present at the direct interface of intercellular host-microbial interactions, and also mediate protein trafficking *via* secreted lectins.

Compared to normal tissues, changes in glycosylation have also been reported under inflammatory stress and cancer progression (*8* - *10*). Unlike the more linear genome and proteome, carbohydrate structures can be linked in a variety of ways like the stereochemistry of the anomeric position, carbohydrate linkage, and overall sequence of carbohydrates (Figure 1.2). Even with a simple disaccharide, one can envision a possible 8 different structural isomers, thus highlighting the combinatorial diversity of carbohydrates. In terms of types of glycans, there are three main classes of mammalian glycoconjugates, *N*-linked, Ser/Thr-linked, and lipid-linked oligosaccharides (Figure 1.1). Based on the collective knowledge of known glycan structures, it has been recently estimated that there are around 7000 different glycan-binding features in the mammalian glycome (*11*).

Due to the staggering complexities of cellular glycosylation, a practical and general method to analyze these carbohydrates in a high-throughput manner was greatly needed. To address this issue, protein microarrays utilizing carbohydrate-binding proteins (CBPs) were developed. This technology is analogous to DNA microarray technology, in which oligonucleotides are printed in a microscale format. For lectin microarrays, a set of lectins, non-enzymatic CBPs, is arrayed in a nanoscale format, and then tested for activity against fluorescently-labeled glycoproteins or other carbohydrate-containing samples (*12* - *14*). The resultant glycopatterns can be further analyzed to reveal discrete carbohydrate epitopes present in the sample. To overcome inherent pitfalls of native plant lectins, recombinant lectins were created to give our lectin microarrays a distinct advantage over current technologies. Prior to microarray technology, carbohydrate analysis was low-throughput and labor intensive. Methods included whole cell agglutination and mass spectrometry, which required additional purification of the glycan or glycoprotein of interest (*15* - *17*).

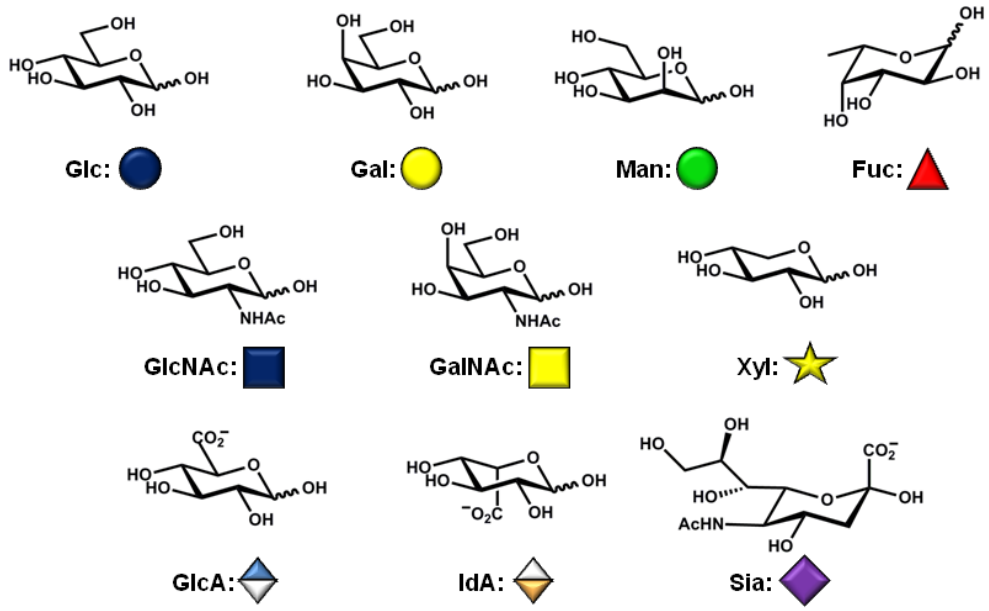

Figure 1.2 The ten major building blocks of the mammalian glycome. These monosaccharides are glucose (Glc), galactose (Gal), mannose (Man), fucose (Fuc), *N*-Acetylglucosamine (GlcNAc), *N*-Acetylgalactosamine (GalNAc), xylose (Xyl), glucuronic acid (GlcA), iduronic acid (IdA), and *N*-Acetylneuraminic acid, or sialic acid (Sia). Symbols used by the Consortium for Functional Glycomics are shown below glycans.

This dissertation focuses on the development of protein microarray technology, specifically, creating new techniques and platforms on which to evolve protein microarrays. The first part of this thesis describes the production of an affinity tag-based method for orienting recombinant lectins. The technique of orienting proteins based on a particular activity tag is not a novel concept, yet no other protein microarray format has been able to address two key issues in site-specific protein orientation. The first issue is that most oriented microarray formats are restricted to the specific affinity tag of interest. This need for a particular affinity handle excludes the printing of non-tagged or differently-tagged proteins, thus limiting the scope of the protein microarray. And second, despite the numerous efforts in protein orientation in microarray technology, not one describes how orientation affects protein activity against a target protein or glycan. The second part of my thesis describes my efforts to modify our protein orientation

strategy to a more simple and systematic method by performing a one-step deposition and orientation of the GST-tagged recombinant lectins. I then demonstrate that this one-pot method can be applied to more complex protein mixtures in order to orient antibodies *in situ*. In applying this technique, I have created a method to expand the current lectin microarray with antibodies, an often under-applied class of CBPs, for direct glycomic profiling. The third section of this dissertation describes the development of binding variants of the current set of recombinant lectins. Due to the ease of genomic manipulation of recombinant proteins, I generated a set of non-binding mutants of the lectins to act as controls in microarray technology. In a concerted effort of generating novel lectins, I will also describe my current efforts in the directed evolution of GafD, a β-GlcNAc-binding lectin derived from *Escherichia coli*. Under the direction of Professor Jonathon Lai (Albert Einstein College of Medicine), I created constructs to produce GafD on the surface of phage, which I show to be a stable, displayed protein capable of binding to a known antigen. I generated a library of GafD variants, yet I had very little luck in selecting for any full-length protein. Although I had little luck with my GafD library in my selections, this does not mean that phage-display technology will not work, but a couple of different suggestions on how to move forward will be discussed. In conclusion, the majority of this dissertation will focus on the inherent advantages of recombinant lectins over traditional, natural plant lectins which are normally used in glycomic analysis.

In the following sections of this chapter, I will discuss the development of protein microarray technology, and then I will specifically discuss the origins and evolution of the lectin microarray platform. I will then conclude this section by discussing the current issues in manufacturing protein microarrays, and I will introduce our approaches to solving these problems.

1.2 Protein Microarray technology

1.2.1 Development of protein microarrays

In the pursuit of the high-throughput analysis of complex samples, protein microarray technology was first developed as a protein-based adaptation of gene microarrays. The seminal work by MacBeath and Schreiber showed that protein microarrays could be used to analyze protein-protein interactions, kinase substrates, and protein-small molecule interactions (*18*). In addition, the pair printed a single protein, FKBP12-rapamycin binding protein (FRP), in the presence of over 10,000 spots of a different protein. When probed with fluorescently-labeled FKBP12, the only interaction observed was the FRB-FKBP12 complex (*18*). This work highlighted that not only can arrayed proteins maintain their function, but it also shows that a vast array of proteins can be miniaturized on a small scale (Figure 1.3). Another textbook case in protein microarray technology was the global analysis of protein-protein interactions of the yeast proteome (*19*). The group cloned 5800 yeast proteins and arrayed them on functionalized glass slides. After probing with calmodulin and phospholipids, the group identified novel binding motifs for both epitopes (*19*). This work highlights the importance that high-throughput analysis of protein-protein interactions which can help identify proteins of unknown functions. The authors also demonstrate the use of site-specific immobilization technique which allows uniform deposition of the entire proteome. All of the proteins were expressed as fusions with a hexahistidine (His_6)-tag and printed on a Nickel^{2+}-nitriloacetic acid (Ni^{2+}-NTA) modified slide surface. This was a good idea, but the group never demonstrated that protein immobilization affects binding activity.

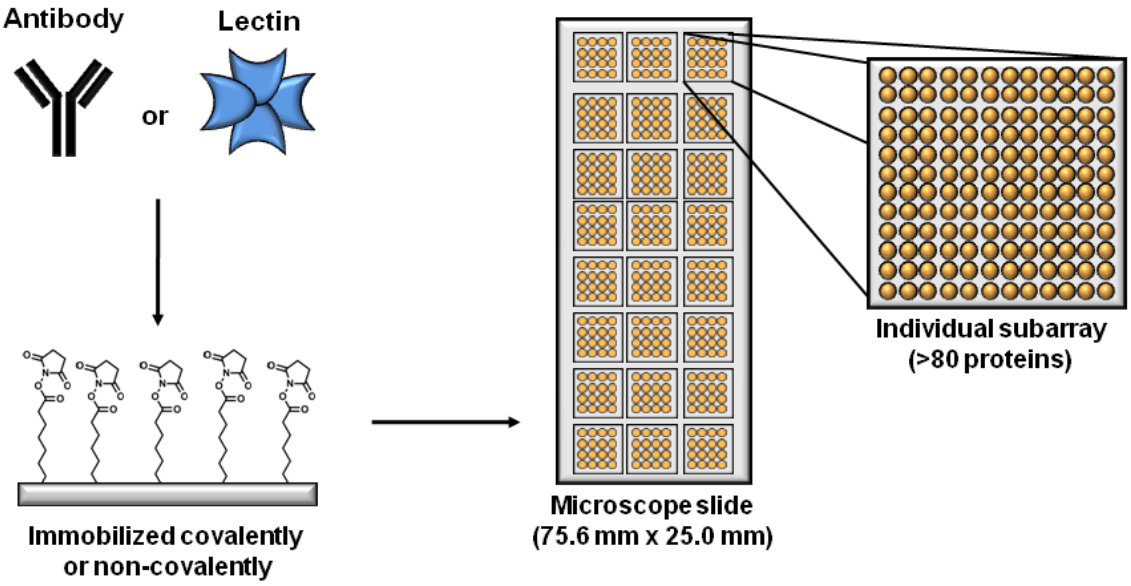

Figure 1.3 Overview of protein microarray technology. Proteins, such as antibodies or lectins, are immobilized either covalently or non-covalently on a microscopy-sized slide. The number of proteins to be arrayed is determined by both the number of subarrays and spot size.

1.2.2 Lectin microarrays for glycomic analysis

To address the issue of global carbohydrate analysis, our lab developed lectin microarray technology, which is the direct immobilization of plant-derived lectins onto a solid support (Figure 1.4). In this technique, fluorescently-labeled glycoconjugates are hybridized to an array of lectins with defined specificities (*20*). Our lab's initial work focused on the discernment of glycopatterns from a couple of different purified glycoproteins (*13*). In this work, Pilobello et al. immobilized 9 well-characterized lectins onto amine-reactive microscope slides. After probing with labeled substrates, the glyocpatterns observed from the lectin array matched to the known glycan epitopes of each respective glycoprotein (*13*).

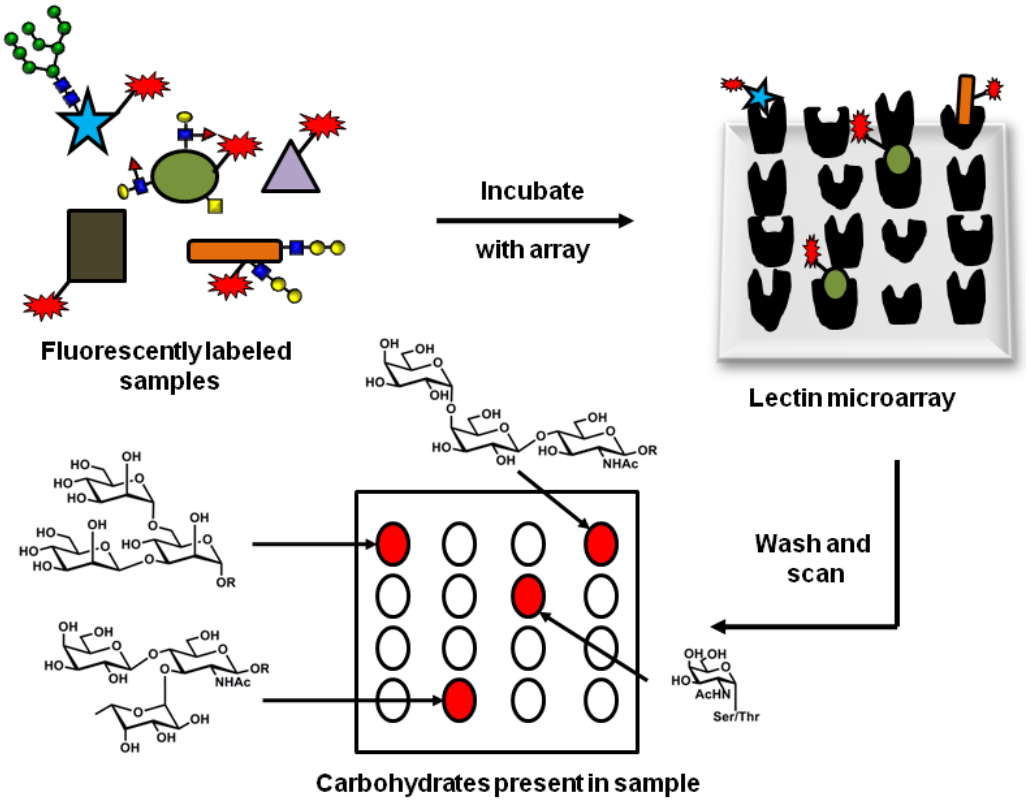

Figure 1.4 Schematic of a lectin microarray experiment. Fluorescently-labeled samples are incubated with a pre-fabricated lectin microarray. After an incubation time, the arrays are washed and scanned at the appropriate wavelength. Known lectin acitivities indicate which glycans are present in the biological sample.

Our group also demonstrated that a lectin microarray could monitor changes in bacterial glycosylation over time (*21, 22*). In this case, Hsu et al. also demonstrated the ability of the lectin microarray to distinguish between pathogenic and non-pathogenic bacteria, which differ in the synthesis of the surface lipopolysaccharide (*21*). Furthermore, in more recent work, Krishnamoorthy et al. examined the biogenesis of the HIV-1 virion (*23*). In theory, HIV-1 replicates by viral budding from the cell surface, which can incorporate host glycoproteins into the secreted microvesicles. Using a two-color labeling scheme (*24*), they demonstrated that the HIV-1 virions and microvesicles share a common glycomic signature that is dependent on the cell line from which they were derived (*23*). Given the previous difficulties in generating a

vaccine against HIV-1, using lectins microarrays highlights these difficulties by analyzing whole cell glycosylation.

Subsequent to our initial work on lectin microarrays, another group, the Hirabayashi laboratory, published a similar platform. Although our two systems are similar overall, the Hirabayashi platform has two distinct differences in their technology: their choice of slide chemistry and the scanning technology (*14*). First, although we initially used epoxy slides to immobilize lectins, we switched to *N*-hydroxysuccinimide (NHS)-activated slides (Nexterion H). These slides are specially designed to both lower overall background signals and maintain protein structure (*25, 26*). The Hirabayashi lab still uses epoxide-derivatized surfaces, which are relatively inexpensive, but lack efficient protein immobilization and low background fluorescence (*13*). Second, our arrays are processed using conventional and widely used DNA array scanners, whereas the Hirabayashi arrays are processed using more expensive evanescent wave scanners (*27*). However, their system is unique since they can detect weaker lectin-glycan interactions, leading to a more sensitive platform (*27*). Despite the differences, their lab demonstrated the differences in glycosylation between non-differentiated and differentiated stem cells using a lectin microarray (*28*). They also showed that the lectin microarray platform can distinguish proteoglycans from healthy and malignant tissues, leading to the discovery of new glycan-biomarkers (*29*).

Current lectin-based microarray technology has enabled the analysis and discovery of new glycoprotein biomarkers and the rapid profiling of cell-surface carbohydrates. With the limited number of plant lectins commercially available, with respect to possible carbohydrate structures, our lab began to search for new lectins with unique binding motifs (*30 - 34*). In both lectin microarray formats, plant lectins have been the sole class of CBPs used, but there are two

major disadvantages to using only plant lectins in an array. First, since the lectins are naturally derived, they often have lot-to-lot variances, further complicating analysis. Second, the majority of commercially available plant lectins are glycosylated. When probed with complex mixtures that contain CBPs, such as bacteria, this increases the occurrence of false-positive signals on the array. In response to these limitations, our lab looked toward lectins derived from bacteria (*35*). Recombinant proteins have an advantage due to the expression and purification of protein without post-translational modifications, allowing for strict quality control. Furthermore, the inclusion of fusion tags can standardize the purification of multiple proteins of varying glycan specificities. Given that bacteria express lectins to mediate host-pathogen interactions, microbial genomes are a rich source of CBPs. Hsu et al. cloned out seven lectins derived from a variety of bacterial sources (Table 1.1). The lectins were optimized for expression and purified *via* the glutathione-*S*-transferase (GST) affinity tag with glutathione (GSH)-sepharose (Figure 1.5). When probed against a glycan microarray developed by the Gildersleeve group (*36, 37*), the lectins retained their known binding preferences (*35*). The recombinant lectins were then immobilized on NHS-activated slides, and probed for activity against glycoproteins and renal and melanoma tumor cell lines from the NCI-60 cell line panel. The resultant glycopatterns were dependent on the probe being tested, and gave distinct glycopatterns. Additionally, the inclusion of a mutant of GafD, GafD-m, confirmed that any binding of wild-type GafD to the sample was carbohydrate-based (*35*). Prior to this work, recombinant lectins have not been utilized in lectin microarrays, and this work highlighted the potential benefits of recombinant lectins over naturally-purified lectins.

Lectin	Source	Binding Specificity
GafD	F17 fimbrae (*Escherichia coli*)	β-GlcNAc
PA-IL	Non-fimbrae (*Pseudomonas aeruginosa*)	Galactose
PA-IIL	Non-fimbrae(*Pseudomonas aeruginosa)*	Fucose/Mannose
PapGII	P-pili (*Escherichia coli*)	GbO4
PapGIII	P-pili (*Escherichia coli*)	GbO5
RS-IIL	Non-fimbrae (*Ralstonia solanacearum*)	Mannose/Fucose
GafD-m	F17 fimbrae (*Escherichia coli*)	~80% reduction in binding

Table 1.1 List of recombinant lectins. This list was genearated by Hsu et al. (*35*)

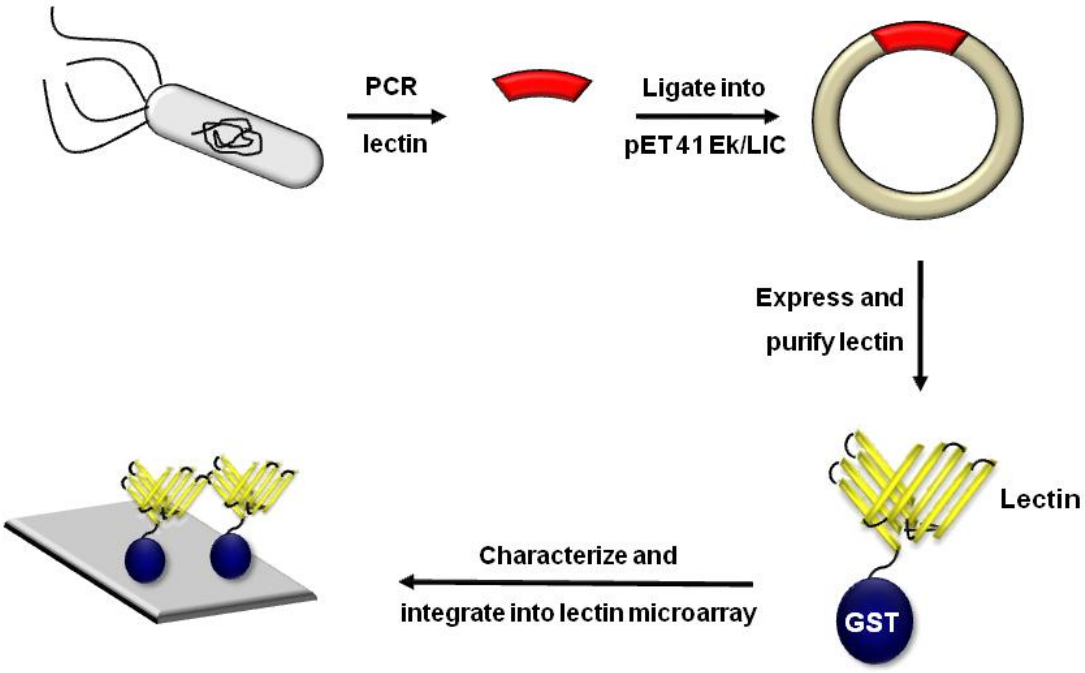

Figure 1.5 Schematic of a standardized method to producing recombinant lectins. Desired lectin gene is PCR amplified and cloned into pET41 vector. The lectin is then expressed and purified *via* the GST-tag. Characterization of glycan-binding activity can be assessed by the detection of multiple fusion tags. Once assayed, the lectin can be incorporated into the lectin microarray, expanding the current set of lectin.

1.2.3 Issues in current microarray fabrication

Like all new technologies, improvements in protein microarray technology are necessary toward widespread implementation. In the production of these arrays, the most important factor is the deposition and the activity of the immobilized protein. One commonly used deposition method is the immobilization *via* side chain lysine amines and amine-reactive functional groups (*38*). However, regardless of the slide surface, most protein microarrays result in immobilization through random deposition onto the slide. Unfortunately, these techniques can deposit a protein in which the functional domain is not oriented toward the protein-sample interface, resulting in diminished activity. This issue prompted the development of various protein immobilization techniques based on known and specific biological interactions, such as biotin-avidin and GST-GSH (*39*). For example, Chen et al. describe an immobilization technique where a boronic acid modified slide binds to the oligosaccharide present on an Fc-tagged lectin (*40*). And since the group modifies the entire slide with the boronic acid, only Fc-tagged proteins can be printed on their array. Although technically unique, the method suffers from the same complicated process and protein exclusivity that every other site-specific immobilization technique does. We believed that we could use the GST domain to immobilize the recombinant lectins on a GSH-modified surface. With our initial success of immobilizing these lectins on a slide treated with GSH, we realized that a GSH-surface would greatly restrict the number of CBPs we could print to the recombinant lectins. By restricting the type of protein to be deposited, we restrict the number of oligosaccharides observed. We wanted to include the non-GST-tagged plant lectins while orienting our GST-tagged recombinant lectins, so we created a dual-surface array. This

platform would maintain the NHS-activated surface for the plant lectins and then a GSH-modified surface for the recombinant lectins (*41*).

In addition to our dual-surface array format, we created a protein orientation method which would site-specifically immobilize our recombinant lectins in a one-step technique (*42*). By supplementing the print buffer with an excess amount of GSH, we can out-compete the lysine side chains for available NHS-activated esters, thereby generating a localized GSH-modified surface that can then orient the recombinant lectins *in situ*. Once we simplified our original orientation technique, we next wanted to expand the technology to other types of CBPs. Carbohydrate-binding antibodies are an often under-utilized CBP in glycomic analysis. In particular, lectin microarrays do not include this class of CBPs, but antibodies can be a good supplement due to their higher specificities compared to lectin-glycan interactions. Antibody microarrays are typically printed onto nitrocellulose-coated slides, but these slides are not compatible with our lectin microarrays. We wondered whether it was possible to not only include antibodies in our array but also orient the proteins by supplementing the print buffer with GSH and a GST-tagged antibody binding protein. We were successful in orienting these antibodies *in situ*, thereby creating the first protein microarray with two classes of proteins.

Another issue in lectin microarray technology is the lack of negative controls. In DNA microarrays, in a given array set it is common to have an array that contains a non-binding gene for a specific data pool (*43*). When the array is hybridized with labeled probes, the non-binding gene should not display fluorescence across multiple arrays. In the case of lectin microarrays, the only technique to determine whether or not a lectin-sample interaction is carbohydrate-based is by performing inhibition experiments (*13*). With the use of recombinant lectins, we can quickly generate binding mutants due to the ease of genetic manipulation. By printing non-binding

variants along with the wild-type lectins, we can easily determine whether an observed lectin binding event is carbohydrate-based. We also show that we can switch the binding preferences of the orthologous lectins PA-IIL and RS-IIL by a single point mutation. In the pursuit generating novel lectins, the Hirabayashi lab attempted to evolve the Ricin B chain lectin from a galactose-binding lectin to a sialic acid-binding protein (*44*). The Ricin B variant that they generate does not lose all galactose-binding activity, and as such, they merely modified the activity to bind more epitopes. With our single point mutant of RS-IIL (A22S), we perform a similar switch in binding preference without having to create a whole new evolution system. However, in order to expand the proteins on our array to increase our view of the glycome, we wanted to create an evolution-based method to generate novel lectins. We chose to evolve GafD from a β-GlcNAc-binder to a β-GalNAc-binding lectin, a relatively simple C4 epimer. With the aid and guidance of Professor Jonathan Lai (Albert Einstein College of Medicine) we were able to produce GafD lectin on the surface of phage as a pIII fusion protein. Although we generated a library of phage clones with 10^9 codon diversity, we were unable to find any definitive enrichment against our target glycan. Fortunately, not all selection conditions were tested meaning that future work could select for a decent binding lectin from the library I generated.

With these advances in lectin microarray technology, we hope to have a broader impact in the protein microarray field. The orientation and then *in situ* orientation of our recombinant lectins solves two major issues in protein arrays: orientation, to enhance lectin sensitivity, and protein diversity. We also demonstrate that non-binding variants of the recombinant lectins can act as efficient controls in the lectin microarray. Also, to generate novel lectins to incorporate into our array, we produce phage-displayed lectins in order to select against new binding proteins.

1.3 References

1. Ohtsubo, K., and Marth, J. D. (2006) Glycosylation in cellular mechanisms of health and disease. *Cell 126*, 855 – 867.
2. Sharon, N. (2006) Carbohydrates as future anti-adhesion drugs for infectious diseases. *Biochim. Biophys. Acta 1760*, 527 – 537.
3. Haltiwanger, R. S., and Lowe, J. B. (2004) Role of glycosylation in development. *Annu. Rev. Biochem. 73*, 491 – 537.
4. Dube, D. H., and Bertozzi, C. R. (2005) Glycans in cancer and inflammation – potential for therapeutics and diagnosis. *Nat. Rev. Drug Discov. 4*, 477 – 488.
5. Marth, J. D., and Grewal, P. K. (2008) Mammalian glycosylation immunity. *Nat. Rev. Immunol. 8*, 874 – 887.
6. Apweiler, R., Hermjakob, H, and Sharon, N. (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta 1473*, 4 – 8.
7. Freeze, H. H. (2006) Genetic defects in the human glycome. *Nat. Rev. Genet. 7*, 537 – 551.
8. Fernandes, B., Sagman, U., Auger, M., Demetrio, M., and Dennis, J. W. (1991) β1-6 Branched oligosaccharides as a marker of tumor progression in human breast and colon neoplasia. *Cancer Res. 51*, 718 – 723.
9. Bos, P. D., Zhang, X. H., Nadal, C., Shu, W., Gomis, R. R., Nguyen, D. X., Minn, A. J., van de Vijver, M. J., Gerald, W. L., Foekens, J. A., and Massague, J. Genes that mediate breast cancaer metastasis to the brain. *Nature 459*, 1005 – 1009.

10. Gu, J., Sato, Y., Kariya, Y., Isaji, T., Taniguchi, N., and Fukuda, T. (2009) A mutual regulation between cell-cell adhesion and N-glycosylation: Implications of the bisecting GlcNAc for biological functions. *J. Proteome Res. 8*, 431 – 435.

11. Cummings, R. (2009) The repertoire of glycan determinants in the human glycome. *Mol. BioSys. 5*, 1087 – 1104.

12. Krishnamoorthy, L., and Mahal, L. K. (2009) Glycomic analysis: an array of technologies. *ACS Chem. Biol. 4*, 715 – 732.

13. Pilobello, K. T., Krishnamoorthy, L., Slawek, D. and Mahal, L. K. (2005) Development of a lectin microarray for the rapid analysis of protein glycopatterns. *ChemBiochem 6*, 985 – 989.

14. Tateno, H., Uchiyama, N., Kuno, A., Togayachi, A., Sato, T., Narimatsu, H., and Hirabayashi, J. (2007) A novel strategy for mammalian cell surface glycome profiling using lectin microarray. *Glycobiology 17*, 1138 – 1146.

15. Pilobello, K. T., and Mahal, L. K. (2007) Deciphering the glycocode: the complexity and analytical challenge of glycomics. *Curr. Opin. Chem. Biol. 11*, 300 – 305.

16. Mahal, L. K. (2008) Glycomics: towards bioinformatic approaches to understanding glycosylation. *Anticancer Agents Med. Chem. 8*, 37 – 51.

17. Rakus, J. F., and Mahal, L. K. (2011) New technologies for glycomic analysis: Toward a systematic understanding of the glycome. *Annu. Rev. Anal. Chem. 4*, 367 – 392.

18. MacBeath, G., and Schreiber, S. L. (2000) Printing proteins as microarrays for high-throughput function determination. *Science 289*, 1760 – 1763.

19. Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bildingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R. A., Gerstein, M., and

Snyder, M. (2001) Global analysis of protein activities using proteome chips. *Science 293*, 2101 – 2105.

20. Pilobello, K. T., and Mahal, L. K. (2007) Lectin microarrays for glycoprotein analysis. *Meth. Mol. Biol. 385*, 193 – 203.

21. Hsu, K.-L., Pilobello, K. T., and Mahal, L. K. (2006) Analyzing the dynamic bacterial glycome with a lectin microarray approach. *Nat. Chem. Biol. 2*, 153 – 157.

22. Hsu, K.-L., and Mahal, L. K. (2006) A lectin microarray approach for the rapid analysis of bacterial glycans. *Nat. Protoc. 1*, 543 – 549.

23. Krishnamoorthy, L., Bess, J. W. Jr., Preston, A. B., Nagashima, K., and Mahal, L. K. (2009) HIV-1 and microvesicles from T cells share a common glycome, arguing for a common origin. *Nat. Chem. Biol. 5*, 244 – 250.

24. Pilobello, K. T., Slawek, D. E., and Mahal, L. K. (2007) A ratiometric lectin microarray approach to analysis of the dynamic mammalian glycome. *Proc. Natl. Acad. Sci. 104*, 11534 – 11539.

25. Fernandez, I. C. S., van der Mei, H. C., Lochhead, M. J., Grainger, D. W., and Busscher, H. J. (2007) The inhibition of the adhesion of clinically isolated bacterial strains on multi-component cross-linked poly(ethylene glycol)-based polymer coatings. *Biomaterials 28*, 4105 – 4112.

26. Harbers, G. M., Emoto, K., Greef, C., Metzger, S. W., Woodward, H. N., Mascali, J. J., Grainger, D. W., and Lochhead, M. J. (2007) Functionalized Poly(ethylene glycol)-based bioassay surface chemistry that facilitates bio-immobilization and inhibits nonspecific protein, bacterial, and mammalian cell adhesion. *Chem. Mater. 19*, 4405 – 4414.

27. Uchiyama, N., Kuno, A., Tateno, H., Kubo, Y., Mizuno, M., Noguchi, M., and Hirabayashi, J. (2008) Optimization of evanescent-field fluorescence-assisted lectin microarray for high-sensitivity detection of monovalent oligosaccharides and glycoproteins, *Proteomics 8*, 3042 – 3050.

28. Toyoda, M., Yamazaki-Inoue, M., Itakura, Y., Kuno, A., Ogawa, T., Yamada, M., Akutsu, H., Takahashi, Y., Kanzai, S., Narimatsu, H., Hirabayashi, J., and Umezawa, A. (2011) Lectin microarray analysis of pluripotent and multipotent stem cells. *Genes Cells 16*, 1 – 11.

29. Kuno, A., Kato, Y., Matsuda, A., Kaneko, M. K., Ito, H., Amano, K., Chiba, Y., Narimatsu, H., and Hirabayashi, J. (2009) Focused differential glycan analysis with the platform antibody-assisted lectin profiling for glycan-related biomarker verification. *Mol. Cell. Proteomics 8.1*, 99 – 108.

30. Duncan, M. J., Mann, E. L., Cohen, M. S., Ofek, I., Sharon, N., and Abraham, S. N. (2005) The distinct binding specificities exhibited by enterobacterial type 1 fimbriae are determined by their fimbrial shafts. *J. Biol. Chem. 280*, 37707 – 37716.

31. Westerlund, B., Van Die, I., Hoekstra, W., Virkola, R., and Korhonen, T. K. (1993) P fimbriae of uropathogenic Escherichia coli as multifunctional adherence organelles. *Zentrabl. Bakteriol. 278*, 229 – 237.

32. Sharon, N., and Ofek, I. (2000) Safe as mother's milk: carbohydrates as future anti-adhesion drugs for bacterial diseases. *Glycoconj. J. 17*, 659 – 664.

33. Garber, N., Glick, J., Gilboa-Garber, N., and Heller, A. (1981) Interactions of Pseudomonas aeruginosa lectins with Escherichia coli strains bearing blood group determinants. *J. Gen. Microbiol. 123*, 359 – 363.

34. Zinger-Yosovich, K., Sudakevitz, D., Imberty, A., Garber, N. C., and Gilboa-Garber, N. (2006) Production and properties of the native Chromobacterium violaceum fucose-binding lectin (CV-IIL) compared to homologous lectins of Pseudomonas aeruginosa (PA-IIL) and Ralstonia solanacearum (RS-IIL). *Microbiology 152*, 457 – 463.

35. Hsu, K.-L., Gildersleeve, J. C., and Mahal, L. K. (2008) A simple strategy for the creation of a recombinant lectin microarray. *Mol. BioSys. 4*, 654 – 662.

36. Manimala, J. C., Li, Z., Jain, A., Vedbrat, S., and Gildersleeve, J. C. (2005) Carbohydrate array analysis of anti-Tn antibodies and lectins reveals unexpected specificities: Implications for diagnostic and vaccine development. *ChemBiochem 6*, 2229 – 2241.

37. Zhang, Y., Li, Q., Rodriguez, L. G., and Gildersleeve, J. C. (2010) An Array-Based Method To Identify Multivalent Inhibitors. *J. Amer. Chem. Soc. 132*, 9653 – 9662.

38. Hsu, K.-L., Pilobello, K., Krishnamoorthy, L., and Mahal, L. K. (2011) Ratriometric lectin microarray analysis of the mammalian cell surface glycome. *Meth. Mol. Biol. 671*, 117 – 131.

39. Rusmini, F., Zhong, Z., and Feijen, J. (2007) Protein immobilization strategies for protein biochips. *Biomacromolecules 8*, 1775 – 1789.

40. Chen, M. L., Adak, A. K., Yeh, N. C., Yang, W. B., Chuang, Y. J., Wong, C. H., Hwang, K. C., Hwu, J. R., Hsieh, S. L., and Lin, C. C. (2008) Fabrication of an oriented Fc-fused lectin microarray through boronate formation. *Angew. Chem. Int. Ed. Engl. 47*, 8627 – 8630.

41. Propheter, D. C., Hsu, K.-L., and Mahal, L. K. (2010) Fabrication of an oriented lectin microarray. *ChemBiochem 11*, 1203 – 1207.

42. Propheter, D. C., and Mahal, L. K. (2011) Orientation of GST-tagged lectins *via in situ* surface modification to create an expanded lectin microarray for glycomic analysis. *Mol. BioSys. 7*, 2114 – 2117.

43. Lee, P. S., and Lee, K. H. (2000) Genomic analysis. *Curr. Opin. Biotech. 11*, 171 – 175.

44. Yabe, R., Suzuki, R., Kuno, A., Fujimoto, Z., Jigami, Y., and Hirabayashi, J. (2007) Tailoring a novel silaic acid-binding lectin from a ricin-B chain-like galactose-binding protein by natural evolution-mimicry. *J. Biochem. 141*, 389 – 399.

Chapter 2: Orientation of recombinant lectins in a microarray format

2.1 Introduction

A primary concern in protein microarray technology is the activity of the deposited protein. The majority of proteins are deposited onto a reactive surface, with random orientation, which can affect the activity of the printed sample. One way to enhance the activity of immobilized samples on a microarray is *via* protein orientation. The orientation of tagged-proteins onto orthogonally-modified surfaces has the potential of increasing the activity of deposited substrates, enhancing the detection capabilities of a microarray. We demonstrate the effectiveness of protein orientation using our GST-tagged recombinant lectins. By creating a dual-surface array, we are able to incorporate non-tagged proteins along with our GST-tagged lectins in the same array, allowing for the first time the direct comparison of oriented versus non-oriented proteins. We show that orientation *via* the GST-domain enhances the activity of the deposited protein.

2.1.1 Common protein deposition methods

Protein deposition onto microarray surfaces can be divided into two categories: covalent and non-covalent coupling chemistries. For non-covalent attachment, the protein is typically adsorbed into the support. For example, most antibody arrays are printed onto nitrocellulose-modified glass slides (*1*). Nitrocellulose slides are used mainly because of their high adsorption properties, resulting in a greater concentration of deposited protein (*2*). Unfortunately, the typical nitrocellulose slide auto-fluoresces, increasing the background binding of the sample which can disrupt the analysis of a given sample. Due to this flaw, there are several manufacturers of specially-coated slides to decrease auto-fluorescence (*2, 3*). Another non-covalent deposition technique requires the use of amine-coated slide surfaces which attract negatively-charged ions.

The use of these slides is typically reserved for DNA microarrays, where the phosphate backbone is an ideal ionic partner to the positively-charged amines, but some groups have used these slides to deposit proteins (*4, 5*). These non-covalent methods of attachment are good for general use in microarray technology, but the deposition is non-covalent and the orientation of the protein is not controlled (*6*).

The covalent attachment of proteins on a solid support is typically performed through the terminal amines on side chain lysines or the sulfhydryl groups on cysteine residues (Figure 2.1). For amine-based coupling chemistry, typical slide surfaces are composed of *N*-hydroxysuccinimide- (NHS)-activated esters (*7 - 9*), aldehydes (7, *10, 11*), and epoxides (7, *10, 12*). Of these three groups, the least used is the aldehyde, which forms a Schiff base with the deposited protein. In our lab's seminal work in lectin microarray technology, Pilobello et al. printed lectins on both aldehyde- and epoxide-modified surfaces and found that the aldehyde slide gave overall higher activity (*10*). Since that time, we have stopped using aldehyde slides, but work by other groups show that these slides are still commonly used (*11*). In terms of lectin microarray technology, epoxide-derivatized surfaces are the slide surface of choice for the Hirabayashi lab, and their associated company GP Biosciences (*12*). In our hands, the epoxide-coated slides gave significant background fluorescence, however the Hirabayashi slides are designed to probe for lectin-glycan interactions in the solution phase using evanescent wave technology (*13*). This technology requires expensive equipment in lieu of a traditional microarray scanner. Epoxy slides are stable under various pH buffers, humidity, and can react with various functional groups, although the primary reactive group is the free amine of lysine (*7*).

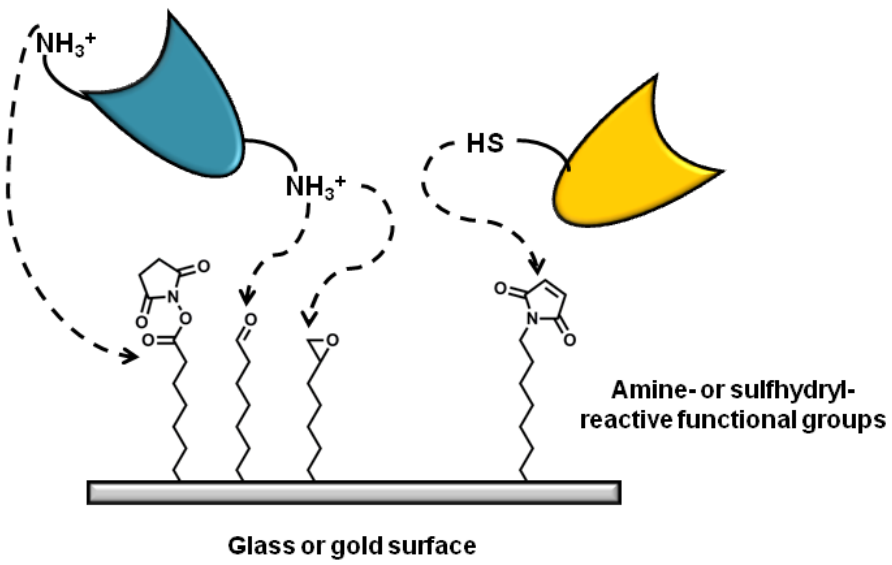

Figure 2.1 Common covalent protein immobilization strategies. The two most common covalent coupling techniques are through amine- and sulfhydryl-functional groups on the deposited protein. Amines from side chain lysine groups can react with NHS-activated esters, aldehydes, and epoxide functional groups. Sulfhydryl-reactive functional groups include maleimide, vinyl sulfones, and various disulfide functional groups.

However, aside from poor background fluorescence, epoxide slides have slow reactivity and, consequently, each protein can be deposited differently depending on incubation times (*14*). As a result, the Mahal lab has been using NHS-activated slides purchased from Schott North America (Nexterion H), technology originally developed by the Grainger laboratory (*15, 16*). These slides are specially developed with NHS-activated esters embedded into a polyethylene glycol (PEG) matrix, creating a 3D polymer coating which can maintain the structural integrity of the deposited protein (*15*). The PEG layer is a highly hydrophilic layer which results in very low background binding from a given sample (*16*). Given the work of our lab wherein our lectin microarrays are probed for activity against complex glycan mixtures such as cellular micellae, whole bacteria, and viral particles, the low background properties of the Nexterion H slide have proven useful in post-hybridization analysis.

Covalent attachment of proteins *via* cysteine residues occurs through selective sulfur-reactive chemistry (Figure 2.1). For example, the most prominently used coupling molecule is maleimide, which serves as an effective Michael acceptor (*17 - 19*). Another substrate used in sulfhydryl-based protein coupling is the Michael addition into vinyl sulfone (*20*). In accordance with the low background of PEG-surfaces, several groups have created hybrid vinyl sulfone/PEG slide surfaces which lower overall background fluorescence (*21*). When the protein is deposited in a buffer below pH 9, these sulfur-reactive functional groups are highly selective for cysteine residues (*20*). This slide coupling method is very selective, yet the availability of solvent-exposed cysteine residues varies from protein to protein, if there are any at all. Although the covalent or non-covalent attachment of proteins to a slide surface is a simple and cost-effective immobilization technique, there is little control on the orientation of the protein acitve site or binding domain, resulting in aberrant activity.

2.1.2 Protein immobilization through affinity tags

Although initially used to facilitate protein production, affinity tags can also be used to track proteins in more complex systems. In the post-genomic era, the high-throughput expression of affinity-tagged systems has become increasingly important in analyzing protein-protein and protein-ligand interactions (*22*). Several groups have attempted to immobilize proteins *via* affinity tags, presumably orienting the protein in a more favorable fashion (*7*). However, no group has ever shown that these immobilization techniques improve activity by comparing the same protein randomly immobilized onto a different slide surface, a point addressed in our work. The three most common affinity tags used for site-specific immobilization in protein microarray systems are streptavidin-biotin (*23 - 27*), hexahistidine (His_6)- Ni^{2+}-NTA (*28 - 32*), and GST-GSH (*33 - 36*) interactions (Figure 2.2).

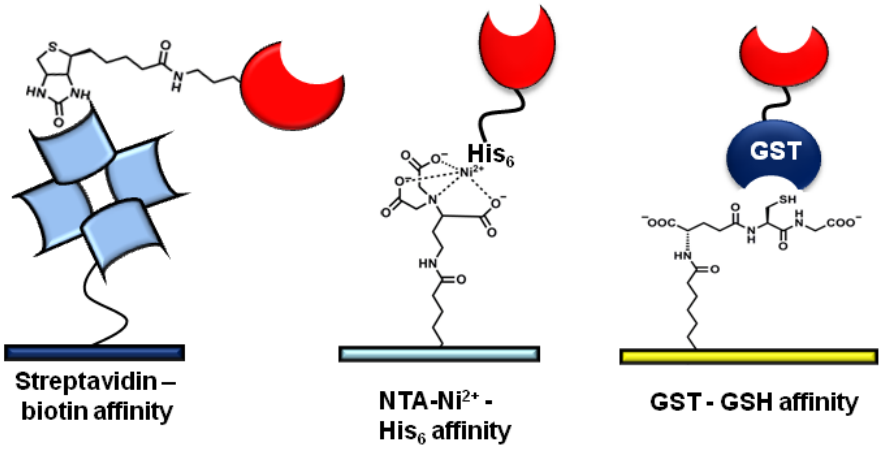

Figure 2.2 Common protein immobilization methods using fusion tags. These interactions include streptavidin-avidin, His_6-Ni^{2+}-NTA, and GST-GSH activities. For streptavidin-biotin interactions, it is believed that one monomer of the streptavidin temtramer is inactive due to inaccessibility on the biotin-binding domain. Therefore, a single streptavidin protein can bind up to three biotinylated proteins. For His_6-Ni^{2+}-NTA interactions, a single molecule of Ni^{2+}-NTA can bind two histidine residues. On the other hand, the GST-GSH interaction forms in a 1:1 molar ratio.

These interactions have been so widely used that these specialized slide surfaces are commercially available. Streptavidin-biotin interactions for use in protein microarray fabrications have become increasingly widespread due to the well-established, and very high affinity interaction between streptavidin and biotin ($K_d \sim 10^{13}$) (*37*). This interaction is heavily used in biotechnology and several groups have generated microarray surfaces capable of immobilizing biotinylated biomolecules. Peluso et al. describe the site-specific immobilization of antibodies onto a streptavidin coated slide (*25*). Their slide surfaces were fabricated with biotin-capped poly-*L*-lysine, a common molecule used in grafting, and the biotin-capped slide was treated with streptavidin, which is a tetrameric protein. And like most streptavidin-biotin microarray systems, the group modified antibodies with a biotin conjugate. The group mildly oxidized the glycans present on the Fc domain; the resulting aldehyde groups were condensed with biotin-hydrazide and applied to the straptavidin-coated slide (*25*). Another group generated

a system to biotinylate recombinant maltose-binding protein (MBP) *in vitro* using an intein-mediated biotinylation scheme and immobilized the substrate onto a streptavidin-coated array (*38*). Although these two groups use an exogenous biotinylation scheme, the use of *in vivo* biotinylation expression systems have the potential to easily generate oriented protein microarrays on streptavidin substrates.

Hexahistidine (His_6) fusion tags are one of the most widely used fusion tags available. When Zhu et al. performed the yeast proteome microarray experiment, the ORFs were cloned into vectors containing *N*-terminal His_6-tags, and were printed onto nickel^{2+}-nitriloacetic acid (Ni^{2+}-NTA)-modified slides (*31*). The His_6-Ni^{2+}-NTA complex has relatively mild affinity (μM) compared to other fusion systems used and is labile under strenuous conditions (*7*). As a consequence, protein immobilization using the standard His_6-tag and Ni^{2+}-NTA may be insufficient to obtain substantial deposited protein (*39*). Given the poor binding affinity of this interaction, several groups have been developing methods to enrich the binding affinity in this system (*32*). For instance, the Piehler group has been experimenting with a synthesized multi-valent NTA ligand that could provide tighter binding to a His_6-tag (*40*). It is a commonly known fact the two histidine side chains interact with one Ni^{2+}-NTA complex. The group developed bi- and tri-dentate NTA ligands to bind stronger to His_6-tagged proteins. Adapting this method, the same group synthesized a biotinylated tri-dentate NTA ligand, which is able to both complex His_6-tagged proteins and immobilize the proteins on a streptavidin-modified slide (*26*). In an analogous method, two other groups have created multi-His-tagged vectors. In one example, Khan et al. cloned two *N*-terminal His_6-tags separated by a spacer of 11 amino acids capable of orienting a GFP protein on a microarray slide (*32*). In another study, Fischer et al. created a similar double His-tagged protein with His_6- and His_{10}-tags. In this study, the authors perform

surface plasmon resonance (SPR) studies on protein immobilization on a Ni^{2+}-NTA surface and demonstrated the labile nature of the His_6-Ni^{2+}-NTA interaction. Although the His_{10}-tag by itself greatly improved binding, the double tagged protein maintained a stable complex, with no visible elution from the slide observed (buffer not containing imidazole) (*39*). Again, these techniques are promising in the development of oriented protein microarrays, but they fail to demonstrate any improvement in activity upon protein orientation.

The third most widely used fusion tag for protein immobilization is glutathione-*S*-transferase (GST), which shows moderate binding activity (μM) to the tri-peptide glutathione (GSH) (*41*). One of the first applications of site-specific immobilizations of GST-tagged proteins was the deposition of GST-calmodulin onto a GSH-treated gold surface (*33*). The group also showed that the GSH-GST interaction was labile and could be eluted off with excess GSH. Although a simple technique, the robust method showed that GSH could be immobilized onto a solid support which could then capture a GST-tagged protein (*33*). Two other techniques of GST-specific protein immobilization used a modified version of a GSH-gold sulfhydryl interaction. One group modified a gold surface with a maleimide-thiol molecule. Once reacted, GSH is added and forms the Michael product, and can then be used to immobilize GST-tagged proteins (*36*). The second group synthesized a dithiol head group and conjugated GSH to it, followed by conjugation to a gold surface. Although unique, the method is not practical given the synthesis of the GSH molecule (3 steps, 6.5% yield) (*42*). In terms of protein microarray fabrication, the work presented by Kawahashi et al. is the most direct use of profiling protein binding from GST-tagged proteins on a GSH-derivatized surface (*35*). The group grafted a slide with GSH- poly-*L*-lysine, and then printed a cadre of GST-tagged proteins that were synthesized *in vitro* and then the resultant arrays were probed for protein-protein interactions with

fluorescently-labeled conjugates (*35*). These previous approaches to immobilizing GST-tagged proteins on a GSH-modified surface are novel in their own right, but suffer from the same issues of other endeavors in tag-specific protein orientation. Specifically, the various groups do not show that the site-specific immobilization of GST-tagged proteins positively affects the activity of the printed protein. Moreover, the affinity tag immobilization methods described above do not allow for the deposition of non-tagged proteins onto the same slide surface, which drastically limits the diversity of proteins present on a microarray.

2.1.3 Fabrication of a GSH-surface to create an oriented recombinant lectin microarray

Prior to fabricating the arrays, we observed two significant pitfalls of current microarray fabrications for protein orientation. First, all microarray formats for protein orientation are for a single type of fusion tag. For example, only His_6-tagged proteins are printed on a Ni^{2+}-NTA surface and GST-tagged proteins are printed only on a GSH-modified surface. By limiting the set of proteins one can print in a microarray, you decrease the observable profile of protein-protein or protein-ligand interactions. Second, of all the work published on oriented protein microarrays, not one actually addresses the question of how the orientation of a protein will affect its affinity or binding activity. All but one study simply assumes that the fusion-tagged protein will be favorably immobilized upon deposition. The one study that attempts to address this question prints two different proteins, GST-tagged and non-tagged, on separate slide surfaces, GSH and non-GSH, respectively (*42*). However, without directly comparing the deposition of the same protein on both surfaces, the group fails to address this important issue in protein microarray technology. Since the GST-GSH interaction has been used in immobilizing GST-tagged proteins, we wanted to directly determine the effect of protein orientation using our unique set of GST-tagged recombinant lectins (*43*).

Our recombinant lectins are cloned into the pET41 vector (Novagen), a vector that contains three different affinity tags (Figure 2.3A). Starting at the *N*-terminus, the vector expresses a GST domain, followed by a His_6 epitope, then an S·tag peptide, followed by the multiple cloning site (MCS). As a result, all of our lectins are located at the *C*-terminus. Fortunately, the binding domains of our current set of lectins are not located at the *N*-terminus meaning that the affinity tags shouldn't interfere with glycan-binding (*43*). This vector also contains a thrombin protease site located between the His_6- and S·tag domains (Figure 2.3B). This feature allows us to cleave the lectin from the GST domain wherein we could assess the effects of the GST-tag on protein orientation and subsequent activity.

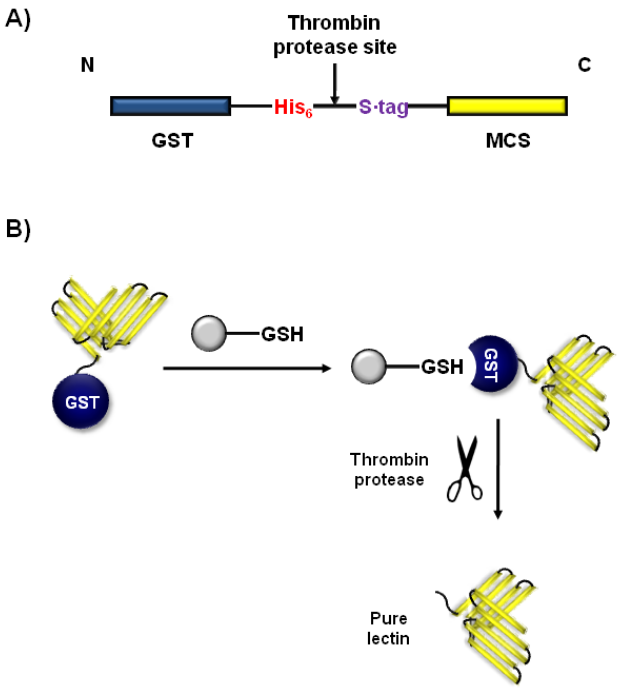

Figure 2.3 Fusion-tag system of the pET41b vector. A) Linear arrangement of the fusion tags and the multiple-cloning site (MCS) of the pET41 vector (Novagen). From the N-terminus, a GST-tag is followed by the hexahistidine tag (His_6), followed by an S·tag domain, followed by the MCS. In between the His_6- and S·tag sites is a thrombin protease domain. An optional C-terminal His_6-tag is available, but in our lectin constructs, we engineer in a stop codon (TAA) at the end of the MCS. B) Schematic of a thrombin-protease treatment of a GST-tagged lectin. Upon immobilization to gluthatione (GSH)-sepharose, the lectin is cleaved from the GST- and His_6- domains by thrombin. Capture of the thrombin protease then yields purified lectin with an *N*-terminal S·tag.

Our traditional lectin microarrays were printed on NHS-activated slides, Nexterion H (Schott North America), on which the activated esters react with side chain lysine groups forming covalent amide bonds, thus immobilizing the lectin to the surface (Figure 2.4A) (*9*). We wondered whether GSH, which contains a free amine, could react with the NHS-esters and form a GSH monolayer (Figure 2.4B). Our initial experiments were geared toward optimizing both the concentration of GSH and buffer composition. We believed that the effectiveness of GSH immobilization is directly dependent on the amount of the small molecule and the pH of the coupling buffer.

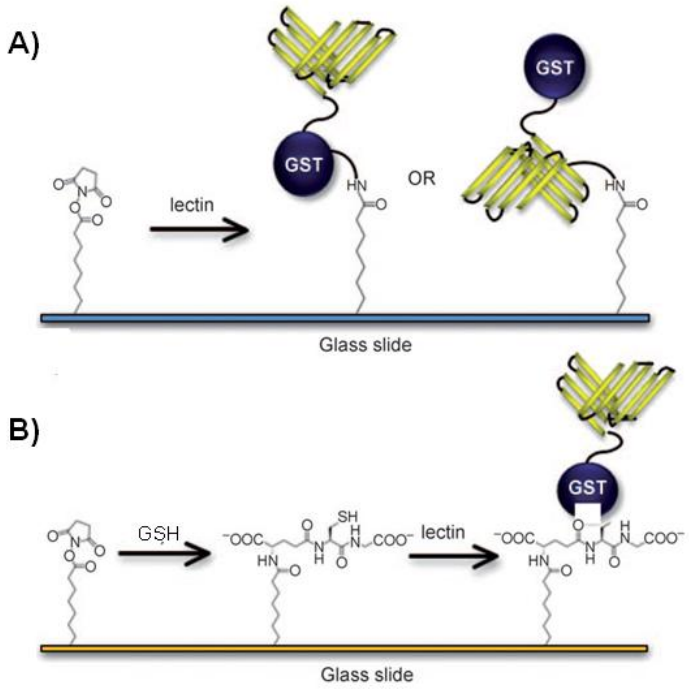

Figure 2.4 Recombinant lectin orientation strategy. A) The standard NHS-activated slides are typically used without further modifications. Printing on the NHS-activated surface results in the random deposition of protein on the surface, perhaps blocking the carbohydrate-binding domain. B) The NHS-activated esters could be reacted with a glutathione (GSH) solution creating a uniform GSH-layer. Then GST-tagged lectins are printed and oriented on the GSH surface (*53*).

We also wanted to print non-GST-tagged plant lectins in the same array in order to expand the range of glycomic markers we can observe. To accomplish this, we decided to divide the arrays in two with a hydrophobic solution using a PAP pen (Beckman Coulter). One half of the array would remain unmodified, maintaining the NHS-surface, and the second half would be reacted with an optimized GSH-solution creating a GSH-surface (Figure 2.5). With this setup, we believed we could print non-tagged lectins on the top half of the array, and then we could directly compare protein deposition and activity by printing the recombinant GST-tagged lectins on both surfaces.

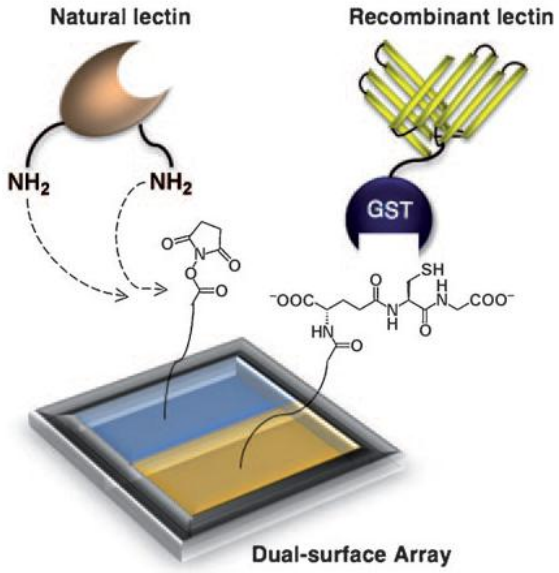

Figure 2.5 Schematic of the dual-surface array. The top half of the array would be left alone while the bottom half of the slide would be reacted with GSH. This would allow the printing of both plant and recombinant lectins on the NHS-activated surface in conjunction with the printing, and hence orientation of our GST-tagged lectins on the bottom half (*53*).

2.2 Results and discussion

2.2.1 Optimization of GSH-immobilization on a solid support

In order to create a dual-surface slide for the direct comparison of oriented and non-oriented lectins, we needed to determine the optimal GSH immobilization conditions. In this work, we did not directly measure GSH deposition as a function of concentration, although we developed one technique for this which will be discussed in Chapter 3. However, we indirectly measured GSH deposition as a function of the binding activity of the deposited protein. For our initial studies, we printed RS-IIL, a GST-tagged mannose-binding lectin derived from the bacterium *Ralstonia solanacearum* (*44*), and probed for activity against Cy5-labeled chicken egg ovalbumin (OVA-Cy5), a known hybrid mannose-containing glycoprotein (*45*). Activity of the immobilized RS-IIL was measured as signal to noise ratio (Absolute fluorescence/background fluorescence, S/N) (Figure 2.6). The coupling of GSH to an NHS-activated ester is an amide-forming reaction, therefore, we reasoned that pH would play an important factor. The primary amine of the tripeptide is the most likely nucleophile. The amide coupling reaction should occur more favorably in a higher pH, in which the primary amine ($pK_a \sim 9$) would not be protonated. Therefore, we tested three common buffers: phosphate buffered saline (PBS, pH 7.4), sodium borate (pH 8.3), and sodium bicarbonate (pH 9.4) containing varying amounts of GSH (200, 100, 50, 25, 12.5, 6.25, and 3.125 mM). RS-IIL was printed at varying concentrations (24, 12, 6, 3, and 1.4 μM) on the GSH-derivatized surface and probed with 600 nM OVA-Cy5. In theory, the GST-GSH binding activity should orient the RS-IIL carbohydrate-binding domain away from the slide surface, making it more accessible to the incoming glycoprotein. Sodium bicarbonate gave us the highest overall binding (Figure 2.6A, paired t-test: bicarbonate vs. PBS, $p = 0.0004$, bicarbonate vs borate, $p = 0.0152$).

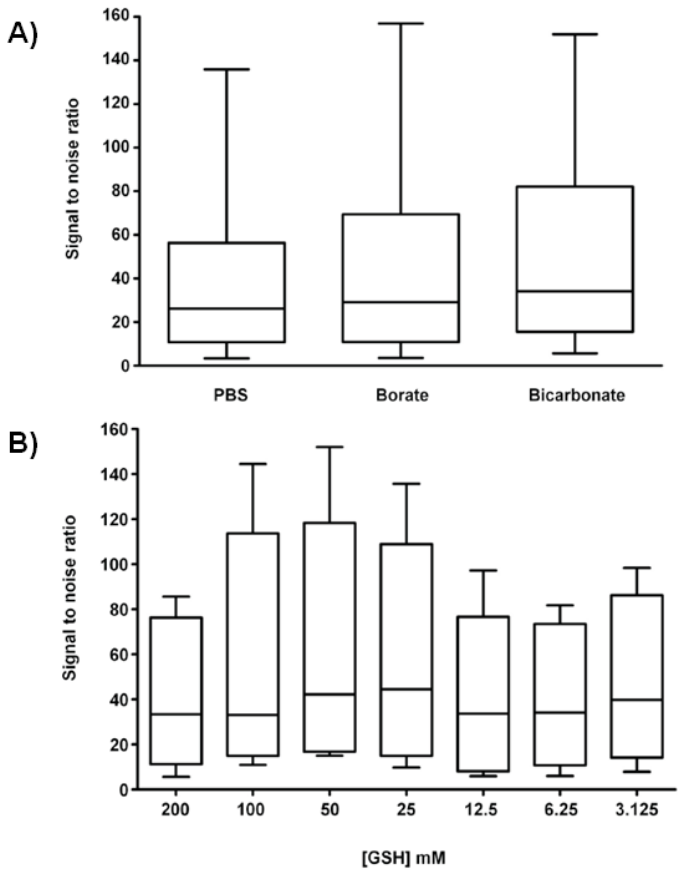

Figure 2.6 Optimization of GSH-deposition. A) Distribution of binding for varying amounts of RS-IIL to OVA-Cy5 with respect to the three different buffers used in the print. Individual arrays were incubated with GSH (200, 100, 50, 25, 12.5, 6.25, 3.125 mM) in either PBS (pH 7.4), sodium borate (pH 8.3), or sodium bicarbonate (pH 9.4) buffers. After processing, RS-IIL was printed on each array (24, 12, 6, 3, and 1.4 μM) and probed with 600 nM OVA-Cy5. B) Distribution of binding for varying amounts of RS-IIL to OVA-Cy5 with respect to different GSH concentrations in sodium bicarbonate buffer. . Individual arrays were incubated with GSH (200, 100, 50, 25, 12.5, 6.25, 3.125 mM) in sodium bicarbonate (pH 9.3) buffer. After processing, RS-IIL was printed in each array (24, 12, 6, 3, and 1.4 μM) and probed with 600 nM OVA-Cy5. Each box plot represents the distribution of binding (S/N) across the different conditions tested. The interquartile range (box) represents the distribution of ~50% of all signals and was used to compare the optimal GSH coupling conditions. Data presented is representative of 3 slides. The optimized buffer conditions were 50 mM GSH in sodium bicarbonate buffer.

For the optimization of GSH concentration, we chose concentrations that centered around the optimized blocking solution recommended by the slide manufacturer, Schott North America. In general after protein printing, the entire slide must be incubated in a solution of 50 mM ethanolamine in sodium borate buffer to react with the remaining NHS-activated esters, thereby inactivating the slide surface. Failure to inactivate the slide results in significant background fluorescence from the primary and/or secondary reactions with fluorescently-labeled probes. Therefore, we surmised that 50 mM GSH would be a good start to test coupling conditions. Comparing the distribution of binding with respect to GSH concentration in all three buffers, we observe that 50 mM GSH is the optimal concentration of GSH (Figure 2.6B). From these initial experiments, we determined that 50 mM GSH in sodium bicarbonate buffer yielded the best activity of our lectin-glycoprotein system.

2.2.2 Activity of oriented versus non-oriented GST-tagged lectins

To assess the effects of orientation on glycan-binding activity, we printed the GST-tagged recombinant lectins on both GSH-modified surfaces (oriented) and on NHS-activated surfaces (non-oriented). Immobilized lectins GafD, a terminal β-GlcNAc-binding lectin derived from *E. coli* (*46*), and RS-IIL were probed for binding activity against varying concentrations of OVA-Cy5 (Figure 2.7). When randomly immobilized onto the NHS-activated surface, non-oriented GafD displayed low levels of glycan-binding. At the highest concentration on OVA-Cy5 (2 μM), observed binding to GafD was very close to our cut-off of a true positive signal (S/N = 5). When oriented on the GSH-modified surface, GafD displayed a remarkable increase in activity (S/N ≈70) against OVA-Cy5 (2 μM), a 17-fold increase in activity compared to the non-oriented GafD (Figure 2.7A and B).

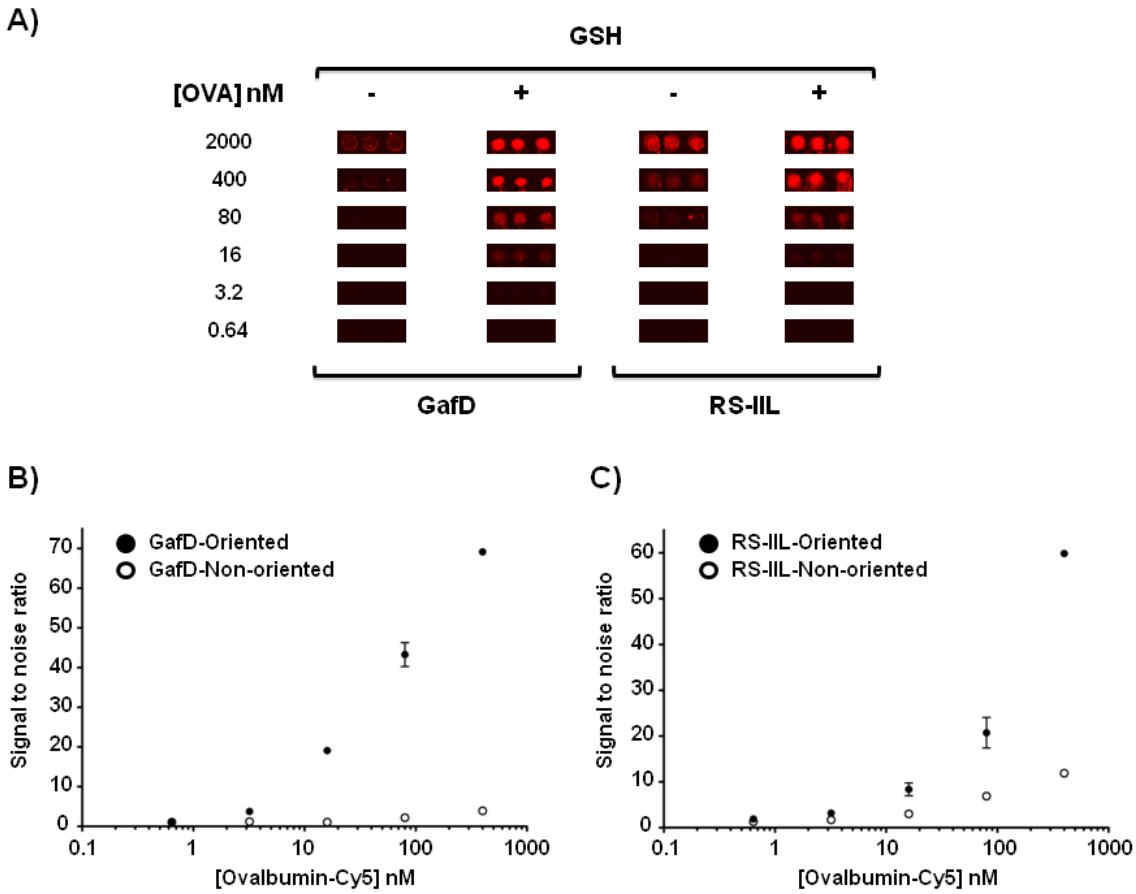

Figure 2.7 Activity of oriented recombinant lectins GafD and RS-IIL against OVA-Cy5. A) Comparison of GafD (10 μM) and RS-IIL (12 μM) binding against decreasing amounts of OVA-Cy5, coupled in either in a random (- GSH, NHS-mediated) or oriented (+ GSH, GSH-mediated) manner. Lectins shown are printed in the same array on the same slide. B) Graphical representation of data shown in A) for GafD as a function of signal to noise. C) Graphical representation of data shown in A) for RS-IIL as a function of signal to noise. All data shown are representative of triplicate slides, and error bars are standard deviations from the median signal to noise (*53*).

Similar results were obtained when comparing oriented versus non-oriented RS-IIL probed against OVA-Cy5. At the highest concentration of OVA-Cy5 tested (2 μM), oriented RS-IIL displayed a 5-fold increase in binding activity compared to the randomly deposited RS-IIL. Herein, we define activity as an increase in observed fluorescence for a particular glycoprotein. Overall, oriented RS-IIL displayed a ~3-fold increase in activity compared to non-oriented RS-IIL over the other concentrations of glycoprotein (Figure 2.7A and C).

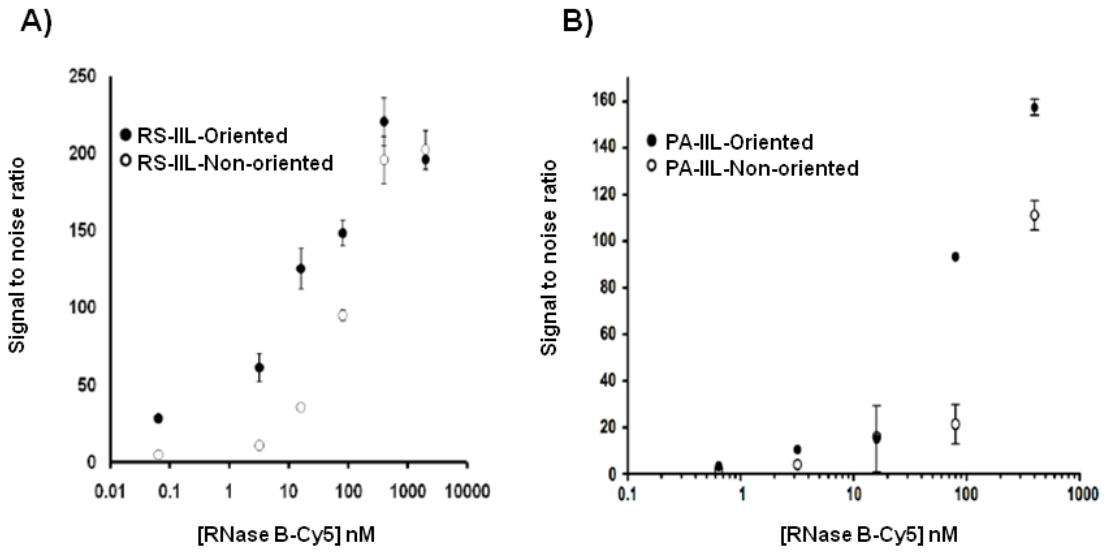

Figure 2.8 Activity of oriented and non-oriented RS-IIL and PA-IIL lectins against Cy5-labeled RNase B (RNase B-Cy5). A) RS-IIL (24 μM) printed on NHS-activated surface (non-oriented) and GSH-modified surface (oriented). B) PA-IIL (12 μM) printed on NHS-activated surface (non-oriented) and GSH-modified surface (oriented). Error bars represent the standard deviation. Data is representative of 3 replicate slides.

We also tested the activity against Cy5-labeled RNase B (RNase B-Cy5), a glycoprotein containing an *N*-linked high mannose site (*51*). As expected, the mannose-binding lectins RS-IIL and PA-IIL, a lectin derived from *Pseudomonas aeruginosa* (*47*), displayed the highest activities with this glycoprotein. With respect to RS-IIL, very little differences exist between the activities of oriented versus non-oriented lectin at higher concentrations of RNase B-Cy5 (Figure 2.8A). At high concentrations of RNase B-Cy5 (2 μM), there is no observable difference in activity between non-oriented and oriented RS-IIL. However, when observing binding at low levels of glycoprotein, a clear pattern emerges showing that oriented RS-IIL shows higher binding sensitivity than randomly deposited RS-IIL. Variations in the level of enhancement of activity of the deposited lectins are expected due to differences in the presentation of the carbohydrate-binding domains, which can skew the apparent binding activity of the adhesin. We believe that lectin orientation through GST-GSH activity results in a more accessible lectin domain. Also, the

sensitivity of detection for the RS-IIL-RNase B-Cy5 interaction was lowered from 3.2 nM to 0.64 nM, a 5-fold increase in detection limits (Figure 2.8A). Although traditionally a fucose-binding lectin, PA-IIL does bind well to high-mannose structures (*48*). The activity of oriented versus non-oriented PA-IIL is ~2-fold over multiple concentrations of RNase B-Cy5 (Figure 2.8B). Compared to the increases in activity observed for GafD and RS-IIL with their respective ligands, PA-IIL displays only a moderate increase in activity upon orientation. Again, the relative differences in activity between the lectins are dependent on the presentation of the glycan epitope.

Overall, we observed increases in activity for the rest of the recombinant lectins against either or both of the Cy5-labeled glycoproteins (Figure 2.9). For example, in the case a PA-IL, a α-galactose-binding lectin also derived from *Pseudomonas aeruginosa* (*49*), did not show an increase in activity when oriented against OVA-Cy5, but it did show a moderate increase with RNase B-Cy5 upon lectin orientation. Interestingly, the PapG lectins (II and III), which are lectins derived from *E. coli* that bind to the galactosylated/GalNAcylated globoside lipids (*50*), showed an increase in activity upon orientation when probed against the two glycoproteins (Figure 2.9A). In the case of these three lectins, it is not surprising to observe an increase in activity against OVA-Cy5, given that the predominant epitope is a hybrid *N*-linked glycan, possibly containing terminal galactose residues (*45*). When it comes to RNase B, it is very surprising to see these lectins, along with GafD, displaying an increase in activity against this high-mannose ligand upon orientation (Figure 2.9B). This glycan-binding activity may be due to contamination of other glycoproteins, or with the ortholog RNase C, a hybrid *N*-linked-containing protein (*51*). Since all of our lectins showed enhanced activity upon orientation, we

needed to determine whether this increase in activity is due to orientation, and not simply because of increased protein deposition.

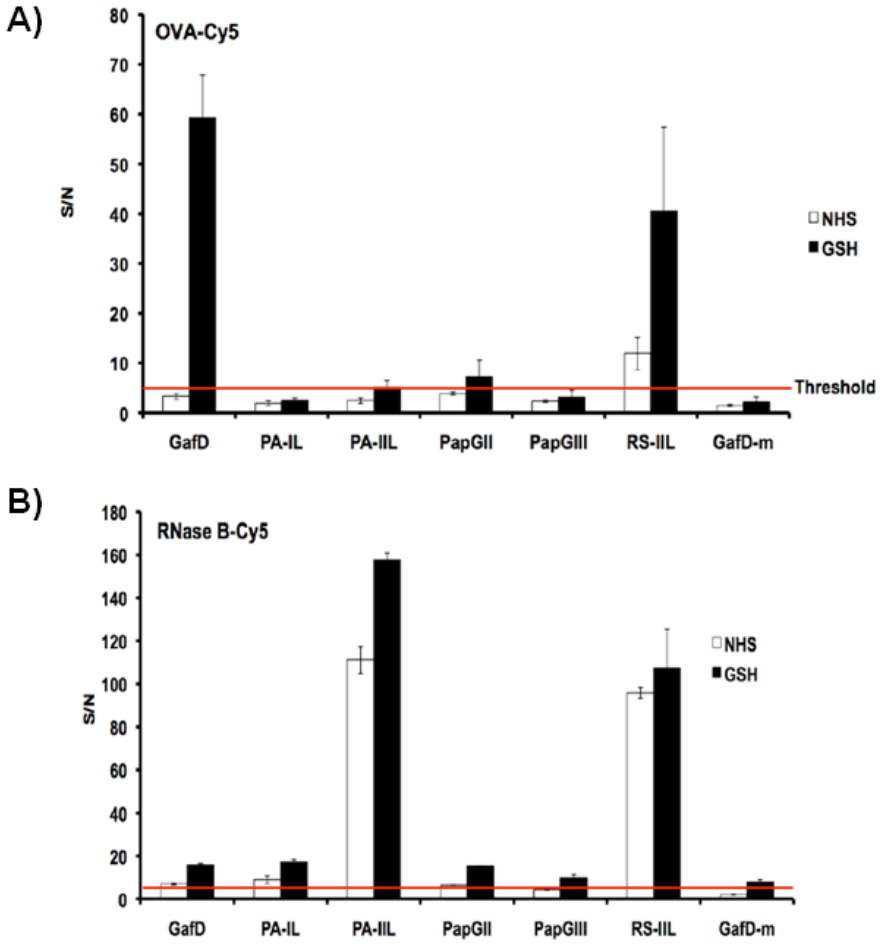

Figure 2.9 Orientation does not affect lectin specificity. All oriented (GSH, black bars) recombinant lectins displayed increases in activity compared to non-oriented lectins (NHS, white bars) against OVA-Cy5 and RNase B-Cy3. A) Graphical representation of signal to noise (S/N) data on dual-surface array probed with OVA-Cy5 (400 nM). B) Graphical representation of S/N data on dual-surface array probed with RNase B-Cy3 (400 nM). GST-tagged lectins (GafD, PA-IL, PA-IIL, PapGII, RS-IIL, GafD-m) were printed at 0.5 mg/mL (~9-12 μM, dependent on the lectin). Average S/N from a single array is shown. Our standard threshold for positive signals (S/N = 5) is represented by a red bar. Enhancement of all positive signals was observed on the array. Error bars represent the standard deviation for 3 spots. Data is representative of 3 replicate arrays.

2.2.3 Determination of lectin deposition upon orientation

One possible reason for the observed increase of activity upon orientation is increased deposition of GST-tagged lectin on the GSH-modified surface. To test this, we labeled GafD with Cy5-NHS and deposited the protein on the two surfaces. We observed a 40% decrease of GafD-Cy5 (Figure 2.10A and B) when printed onto the GSH-modified surface compared to random deposition on the NHS-activated surface. Also, we performed a secondary detection experiment by probing with phycoerythrin-conjugated α-S·tag antibody (α-S·tag-PE). Through observing the Cy3 channel, we can indirectly detect the amount of GafD-Cy5 printed on the slide. When printed onto the GSH-modified surface, we observed a ~4.5-fold increase in α-S·tag-PE binding to GafD (Figure 2.10C and D). In the pET-41 vector, the S·tag domain is nearer to the *C*-terminus of the protein (Figure 2.3A), and given the results of the direct labeling experiment, we concluded that upon orientation *via* the *N*-terminal GST tag, the S·tag domain is also more accessible, hence the increase in detection observed in the Cy3 channel.

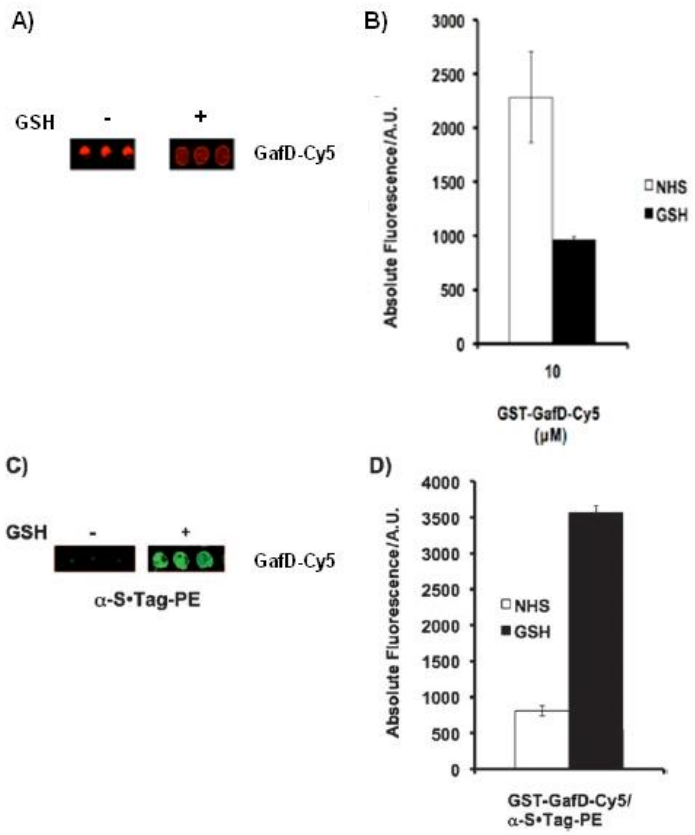

Figure 2.10 Evidence for recombinant lectin orientation. A) Direct detection of GafD-Cy5 (10 μM) printed on the NHS-activated surface (- GSH) and GSH-modified surface (+ GSH). B) Graphical representation of data shown in A). C) Indirect detection of GafD-Cy5 (10 μM) printed on the NHS-activated surface (- GSH) and GSH-modified surface (+ GSH) and detection with α-S·tag-PE. D) Graphical representation of data shown in C). Signals are average between two subarrays with 3 spots per array shown. Error bars represent the standard deviation.

To confirm that orientation of GafD requires the GST-tag, we cleaved the GST-domain with thrombin protease (see Figure 2.3B), and purified the Cy5-labeled, thrombin-cleaved GafD (tc-GafD-Cy5, ~ 25 kDa) (Figure 2.11A). We printed tc-GafD-Cy5 on both the NHS-activated surface and on the GSH-treated surface. Significant protein deposition was observed on the NHS-activated surface, however, only background levels of protein were seen on the GSH-modified surface (Figure 2.11B and C). This experiment validated that the GST-GSH interaction is crucial for protein binding and orientation on the GSH-treated surface.

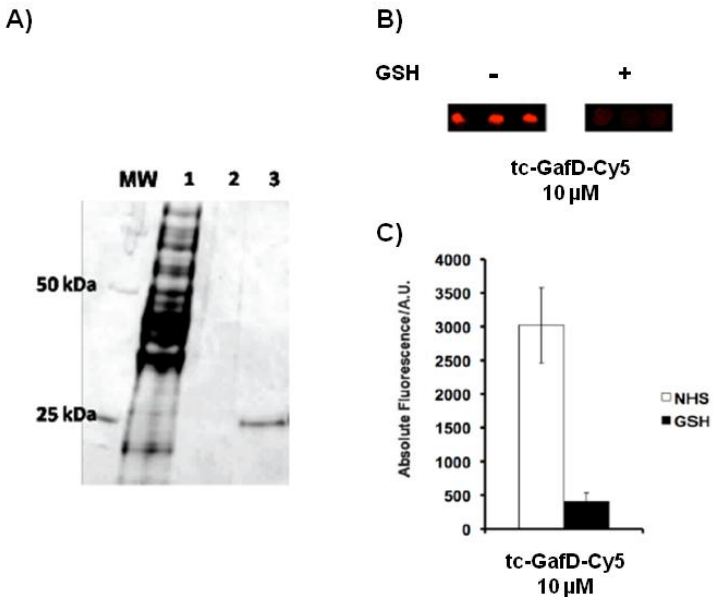

Figure 2.11 Deposition of tc-GafD-Cy5 on the dual-surface array. A) Full-length GafD (~50 kDa) was isolated from a cultured cell pellet (Lane 1) using GSH-sepharose. Immobilized protein was then labeled with Cy5-NHS, then cleaved with thrombin, washed (Lane 2), thus releasing tc-GafD-Cy5 (~23 kDa, lane 3 is supernatant containing free lectin). B) tc-GafD-Cy5 was printed on the NHS-activated and GSH-treated surfaces. C) Graphical representation of data shown in B). Data is representative of 3 arrays. Error bars represent the standard deviation.

2.2.4 Creation of the dual-surface array

To directly compare binding activity of our recombinant lectins, we needed to manufacture a dual-surface array capable of random protein coupling *via* NHS-esters and orientation through GSH-GST activity (Figure 2.12). Considering that plant lectins are the most widely used proteins in glycomic analysis, we also wanted to include these naturally-purified, non-GST-tagged lectins. On the NHS-activated surface (blue outline), we printed 21 naturally-purified plant lectins along with the seven recombinant lectins (with serial dilutions) (See Appendix Table 1 for print list). On the GSH-modified surface (yellow outline), we only printed the GST-tagged recombinant lectins. The dual-surface arrays were then probed with 400 nM OVA-Cy5 (Figure 2.12). As shown below, the glycopattern observed after incubation with the array are in agreement with known glycans present on the glycoprotein (*45*). As expected, we

observed activity against mannose-binding lectins (ConA, RS-IIL, and PA-IIL) and β-GlcNAc-binding lectins (DSA, GS-II, WGA, and GafD).

Figure 2.12 Fabrication of a dual-surface array. Subarrays were generated were subdivided by a hydrophobic solution. A single sub-divided array was probed with 400 nM OVA-Cy5. NHS-activated surface outlined by a blue line, and the GSH-modified surface is outlined in yellow (*53*).

2.3 Conclusions

The creation of a dual-surface array allowed us to have randomly deposited, non-GST-tagged plant lectins printed on an NHS-activated surface in the same array as oriented GST-tagged lectins printed on a GSH-modified surface (*53*). Although we developed this method to be of general use in protein microarray technology, we discovered three main issues with this array platform. First, the production of the dual-surface arrays is very tedious and very prone to human error. Indeed, one minor slip of movement while subdividing the array can completely distort a single subarray and possibly disrupt the printing process. The second issue is the requirement of a GST-tag on the protein deposited. Without subdividing the slide, non-GST-tagged proteins cannot be printed on the GSH-surface. For decades, native, non-tagged, plant lectins have been used to analyze glycosylated substrates (*9, 10, 12, 54*). Lectin microarrays take

advantage of the well-defined set of commercially available lectins. Ideally, these plant lectins would be replaced by recombinant substitutes due to issues mentioned earlier in Chapter 1, however, that is currently not feasible. The requirement for only GST-tagged recombinant lectins on a GSH-modified surface limits the depth of analysis to glycopatterns obtained with only a strict set of recombinant lectins. The third issue is that when probing the GSH-modified surface with complex mixtures, such as cellular micellae (*9*), we observed very high background signal. The coating of the Nexterion H slides was created to maximize protein repulsion, which then dispels non-specific binding interactions, effectively lowering the background levels of fluorescence (*16*). However, when we probed our GSH-modified surface with membrane preparations from some tumor cell lines, we observed significant background fluorescence (Figure 2.13). This is most likely due to the fundamental change of the hydrophobicity and surface charge of the Nexterion H slide.

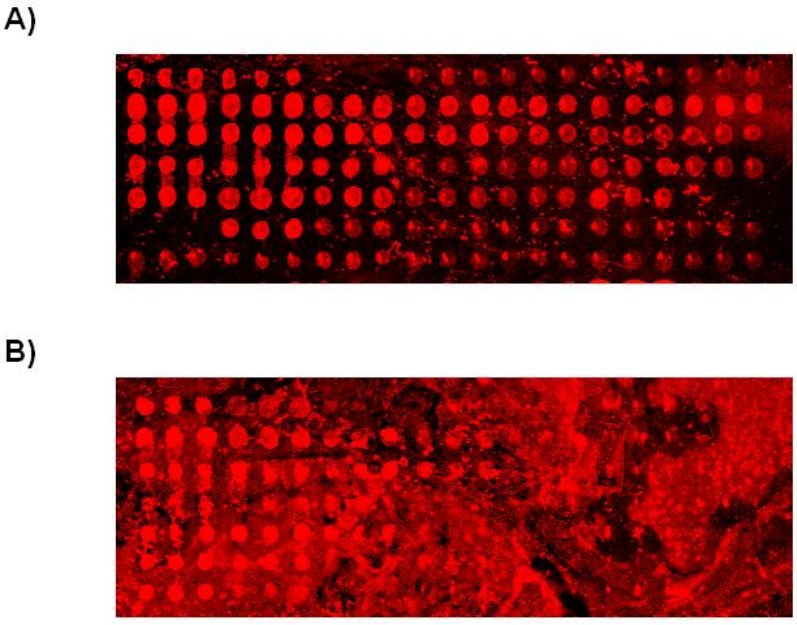

Figure 2.13 GSH disrupts the anti-adhesive coating of the Nexterion H slide. A) Array probed with Cy5-labeled 10 µg of ACHN membrane preps hybridized with a recombinant lectin microarray deposited on an NHS-activated surface. B) Array probed with Cy5-labeled 10 µg of ACHN membrane preps hybridized with a recombinant lectin microarray oriented on a GSH-modified surface.

GSH is a small, zwitterionic peptide, but when immobilized through the primary amine, the *C*-terminal carboxyl group provides a negatively-charged surface, capable of forming ionic interactions with positively-charged samples. Because of these varying limitations, we wanted to develop a method wherein we could preserve the anti-adhesive coating of the Nexterion H slide while taking advantage of the GST-based protein orientation.

Another issue pertains to the construct that we clone our recombinant lectins into. The pET41 vector contains the *N*-terminal GST tag, and as yet, no other company or group to our knowledge has created a *C*-terminal GST vector. It is possible that a GST tag on the *C*-terminus may allow better access to the carbohydrate-binding domain. Also, if one encounters expression difficulties due to truncated products co-purifying on a GSH-column, a *C*-terminal GST-tag would allow the purification of full length GST-tagged proteins. To this end, I modified the pET45b vector, and subcloned in a GST-tag at the *C*-terminus by replacing the S·tag (Figure 2.14). Future experiments in the lab will have to be conducted to prove the utility of this new vector.

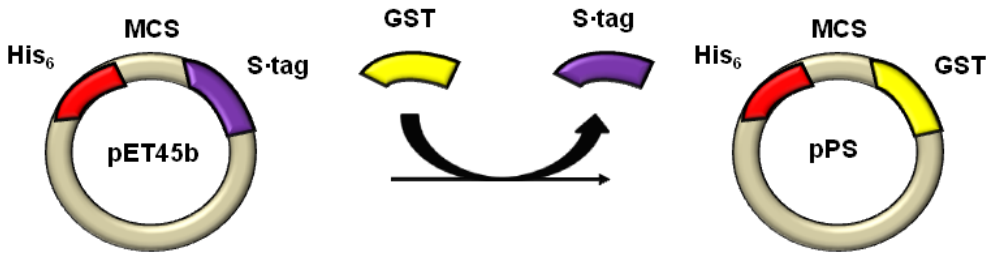

Figure 2.14 Construction of a C-terminal GST-fusion vector, pPS. The S·tag was cut out of the pET45b vector, and a GST construct was ligated into the cut vector yielding pPS.

2.4 Materials and methods

2.4.1 Optimization of glutathione immobilization

Optimization of GSH treatment conditions: *N*-hydroxysuccinimide activated slides (Nexterion slide H, Schott North America) were mounted in a 24-subarray frame (Arrayit Corporation) and buffer or glutathione mix (100 µL) was added to each well. Three buffers were tested: phosphate-buffered saline (PBS, 100 mM sodium phosphate, 150 mM sodium chloride, pH 7.4), sodium borate (50 mM, pH 8.3), and sodium bicarbonate buffer (100 mM, pH 9.4). Different concentrations of glutathione were prepared in these buffers at the following concentrations: 200, 100, 50, 25, 12.5, 6.25, and 3.13 mM. The remaining subarrays were incubated with buffer to act as a control. Slides were incubated with the solutions for 1 h, removed from the frames, and dried on a slide spinner. The slides were washed three times with PBS-T (PBS + 0.005% Tween 20) and once with PBS. Slides were then blocked with ethanolamine (50 mM) in sodium borate buffer for one hour. Slides were then washed and dried as described above, and used for printing. RS-IIL was printed at various concentrations (24, 12, 6, 3, and 1.4 µM) as previously described (*53*). Briefly, lectins were diluted in print buffer (1 mM $CaCl_2$, 1 mM $MgCl_2$, 1 mM of the appropriate monosaccharide (see Appendix Table 1), 0.5 mg/mL BSA, in PBS). The lectin solutions were loaded into a 384-well plate, and the glass slides were printed by using a Spotbot personal microarrayer with SMP3 pins (TeleChem International Inc., Sunnyvale, CA, USA). Three spots per lectin were printed. During the printing, the arrayer was kept between 50–60% humidity and 8°C. Upon completion of the array print, slides were incubated for 2 h, and then blocked and washed as described above. Slides were then placed into the 24-subarray frames and incubated with chicken egg ovalbumin (600 nM, Sigma) fluorescently labeled with Cy5 (Cy5–NHS ester, GE Life Sciences) in PBS-T++ (PBS, 0.005%

Tween 20, 1 mM $CaCl_2$, 1 mM $MgCl_2$) for 2 h at room temperature. Upon completion, the slides were washed and dried as described above and scanned on a Genepix 4100 A slide scanner (Molecular Devices). Genepix Pro 5.1 software (Molecular Devices) was used for extraction of the data. Microsoft Excel and Graphpad Prism 4.0 software were used for statistical analysis and to generate graphs and tables. For data analysis, lectin binding was defined as the signal to noise ratio (S/N): median fluorescence intensity of the sample at 635 nm/median fluorescence intensity of the local background at 635 nm. The average and standard deviation of the S/N for the three replicate spots were reported.

2.4.2 Creation of the dual-surface lectin microarray

To create distinct chambers in a single subarray, a 16-well FAST frame hybridization chamber (Whatman) was coated with a hydrophobic solution from a Super PAP Pen (Beckman-Coulter). The coated FAST frame was then used to stamp the *N*-hydroxysuccinimide activated slides (Nexterion slide H) to create 16 individual subarrays. Each subarray was then further chambered by physically drawing a barrier (~2/3 the length of the subarray) using the hydrophobic solution. To print two distinct lectin panels into a single subarray, the subgrid dimensions and the lateral and vertical offsets must be adjusted accordingly (Figure 2.15).

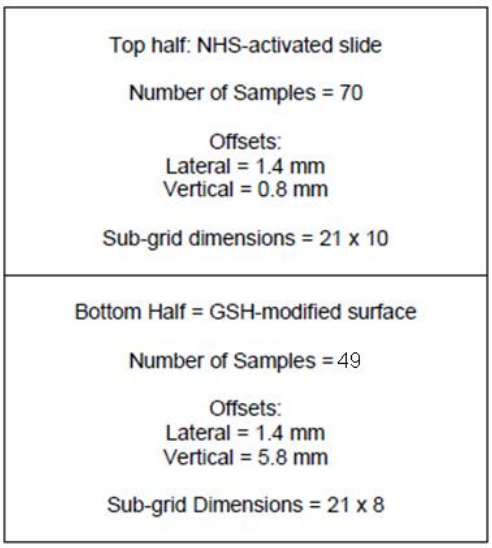

Figure 2.15 Dual-surface lectin microarray layout.

Currently, the maximum numbers of probes with three replicate spots per probe that can be printed in the top and bottom chambers are 70 and 49, respectively (Note: these settings are optimized for the Spotbot Personal Microarrayer and the software provided by Arrayit to print multiple subarrays onto glass slides. Additional optimization would be needed to adapt this configuration to other systems). The full panel of both recombinant and plant lectins (see Appendix A Table 1) was printed onto the top half of each subarray by using the printing procedure described above. Prior to printing, the bottom half of each subarray was incubated with 10 μL of the GSH buffer (50 mM GSH, pH 9.4) and allowed to incubate during the printing procedure. Upon completion of the printing, the slides were incubated for ~2 h to ensure that the coupling process was complete. Please note, total incubations times of longer than 5 h (print time+incubation) can lead to evaporation of the GSH solution. The slides were washed three times with PBS-T and then once with PBS. Slides were then blocked with 50 mM ethanolamine in sodium borate buffer (pH 8.5) for 1 h at room temperature. Slides were washed and dried as described above and were then ready for printing the second panel of lectins (lectins 1–49, see

Appendix A Table 1) in the bottom subarray. Upon completion of the second print, the slides were incubated for ~2 h at room temperature, washed, and dried as described above. Slides were then placed into the 16-well FAST frame hybridization chamber. Each subarray was incubated OVA–Cy5 or RNaseB–Cy5 (New England Biolabs) in PBST++ at varying concentrations for 2 h at room temperature. Upon completion, slides were gently washed and dried, as described previously. Slides were then scanned on a Genepix 4100 A slide scanner, and data was extracted and analyzed as described above.

2.4.3 Cy5-labeling, thrombin treating, and deposition of GafD on a dual-surface array

A single cell pellet, grown from a 1 L culture, was resuspended and lysed as previously described (*8*). The lysate was split into three fractions and loaded onto GSH-sepharose (GE Healthcare, 20μL of beads per mL of lysate). After incubation at room temperature (30 min) the beads were pelleted and washed 3 x with 100 mM $NaHCO_3$ buffer (pH 9.3). The beads were then reacted with 5 μL of 15 μM NHS-Cy5 for 1 hr at room temperature to label the protein. After labeling, GST-GafD-Cy5 was either eluted from the beads using 10 mM GSH in 50 mM Tris*HCl (pH 8.0) or washed 3x with PBS and subjected to thrombin treatment (5 U, 1 hr, RT) to release tc-GafD-Cy5. Thrombin cleaved GafD was treated with 100 μL of benzamidine beads (GE Healthcare) to remove thrombin. The released labeled proteins were then dialysed against PBS, flash frozen and stored at -80 °C until use on the microarrays.

2.4.4 Cloning of pPS

The GST domain was PCR amplified from the pET41b vector containing GafD in the multiple cloning site with the appropriate restriction sites: *XhoI* Forward: 5’ – GAT ATA CAT **CTC GAG** ATG TCC CCT ATA CTA GGT TAT TGG – 3’; *PacI* Reverse: 5’ – CCA TCC GAT T**AA TTA ATT** TTG GAG GAT GGT CGC CAC CAC C – 3’. PCR reactions were

performed with the Hot Start DNA Polymerase (New England Biolabs (NEB), #F120S) and reactions were tailored to product specifications. The PCR product and pET45b DNA were treated with both *XhoI* (NEB, #R0146) and *PacI* (NEB, #R0547) in buffer 4 (NEB). After PCR purification (Qiagen, #28104), the products were mixed and ligated with T4 DNA ligase (NEB, #M0202) and incubated overnight at room temperature. After another PCR purification, the eluant was transformed into electrocompetent NovaBlue Gigasingles (Novagen, #71227), grown, and purified DNA was isolated using the Qiaprep Mini Kit (Qiagen, #27106), and sequenced.

2.5 References

1. Haab, B. B. (2006) Applications of antibody array platforms. *Curr. Opin. Biotech. 17*, 415 – 421.
2. Hucknall, A., Kim, D.-H., Rangarajan, S., Hill, R. T., Reichert, W. M., and Chilkoti, A. (2009) Simple fabrication of antibody microarrays on nonfouling polymer brushes with femtomolar sensitivity for protein analytes in serum and blood. *Adv. Mat. 21*, 1968 – 1971.
3. Wingren, C., and Borrebaeck, C. A. K. (2007) Antibody-based mircorarrays: From focused assays to proteome-scale analysis. *Chem. Mat. Sci. Bioarrays Pt. III,* 175 – 189.
4. Kannan, B., Castelino, K., Chen, F. F., and Majumdar, A. (2006) Lithographic techniques and surface chemistries for the fabrication of PEG-passivated protein microarrays. *Biosensors Bioelectronics 21*, 1960 – 1967.
5. Patwa, T. H., Wang, Y., Miller, F. R., Goodison, S., Pennathur, S., Barder, T. J., and Lubman, D. M. (2008) A novel phosphoprotein analysis scheme for assessing changes in premalignant and malignant breast cell lines using 2D liquid separations, protein microarrays and tandem mass spectrometry. *Proteomics Clin. Appl. 3*, 51 – 66.
6. Tomizaki, K., Usui, K., and Mihara, H. (2010) Protein-protein interactions and selection: array-based techniques for screening disease-associated biomarkers in predictive/early diagnosis. *FEBS J. 277*, 1996 – 2005.
7. Rusmini, F., Zhong, Z., and Feijen, J. (2007) Protein immobilization strategies for protein biochips. *Biomacromolecules 8*, 1775 – 1789.
8. Propheter, D. C., Hsu, K.-L., and Mahal, L. K. (2011) Recombinant lectin microarrays for glycomic analysis. *Meth. Mol. Biol. 723*, 67 – 77.

9. Hsu, K.-L., Pilobello, K., Krishnamoorthy, L., and Mahal, L. K. (2011) Ratriometric lectin microarray analysis of the mammalian cell surface glycome. *Meth. Mol. Biol. 671*, 117 – 131.

10. Pilobello, K. T., Krishnamoorthy, L., Slawek, D. and Mahal, L. K. (2005) Development of a lectin microarray for the rapid analysis of protein glycopatterns. *ChemBiochem 6*, 985 – 989.

11. Chen, M. L., Adak, A. K., Yeh, N. C., Yang, W. B., Chuang, Y. J., Wong, C. H., Hwang, K. C., Hwu, J. R., Hsieh, S. L., and Lin, C. C. (2008) Fabrication of an oriented Fc-fused lectin microarray through boronate formation. *Angew. Chem. Int. Ed. Engl. 47*, 8627 – 8630.

12. Tateno, H., Uchiyama, N., Kuno, A., Togayachi, A., Sato, T., Narimatsu, H., and Hirabayashi, J. (2007) A novel strategy for mammalian cell surface glycome profiling using lectin microarray. *Glycobiology 17*, 1138 – 1146.

13. Uchiyama, N., Kuno, A., Tateno, H., Kubo, Y., Mizuno, M., Noguchi, M., and Hirabayashi, J. (2008) Optimization of evanescent-field fluorescence-assisted lectin microarray for high-sensitivity detection of monovalent oligosaccharides and glycoproteins, *Proteomics 8*, 3042 – 3050.

14. Mateo, C., Abian, O., Fernandez-Lorente, G., Pedroche, J., Fernandez-Lafuente, R., Guisan, J. M., Tam, A., and Daminati, M. (2002) Epoxy sepabeads: a novel epoxy support for stabilization of industrial enzymes via very intense multipoint covalent attachment. *Biotech. Prog. 18*, 629 – 634.

15. Fernandez, I. C. S., van der Mei, H. C., Lochhead, M. J., Grainger, D. W., and Busscher, H. J. (2007) The inhibition of the adhesion of clinically isolated bacterial strains on multi-

component cross-linked poly(ethylene glycol)-based polymer coatings. *Biomaterials 28*, 4105 – 4112.

16. Harbers, G. M., Emoto, K., Greef, C., Metzger, S. W., Woodward, H. N., Mascali, J. J., Grainger, D. W., and Lochhead, M. J. (2007) Functionalized Poly(ethylene glycol)-based bioassay surface chemistry that facilitates bio-immobilization and inhibits nonspecific protein, bacterial, and mammalian cell adhesion. *Chem. Mater. 19*, 4405 – 4414.

17. Viitala, T., Vikholm, I., and Peltonen, J. (2000) Protein immobilization to a partially cross-linked organic monolayer. *Langmuir 16*, 4953 – 4961.

18. Gauvrea, V., Chevallier, P., Vallieres, K., Petitclerc, E., Gaudreault, R. C., and Laroche, G. (2004) Engineering surfaces for bioconjugation: developing strategies and quantifying the extent of the reactions. *Bioconjug. Chem. 15*, 1146 – 1156.

19. Jongsma, M. A., and Litjens, R. H. (2006) Self-assembling protein arrays on DNA chips by auto-labeling fusion proteins with a single DNA address. *Proteomics 6*, 2650 – 2655.

20. Masri, M. S., and Friedman, M. (1988) Protein reactions with methyl and ethyl vinyl sulfones. *J. Protein Chem. 7*, 49 – 54.

21. Rizzi, S. C., and Hubbell, J. A. (2005) Recombinant protein-co-PEG networks as cell-adhesive and proteolytically degradable hydrogel matrixes. Part I: Development and physicochemical characteristics. *Biomacromolecules 6*, 1226 – 1238.

22. Xie, H., Guo, X.-M., and Chen, H. (2009) Making the most of fusion tags technology in structural characterization of membrane proteins. *Mol. Biotech. 42*, 135 – 145.

23. Pavlickova, P., Knappik, A., Kambhampati, D., Ortigo, F., and Hug, H. (2003) Microarray of recombinant antibodies using a streptavidin sensor surface self-assembled onto a gold layer. *BioTechniques 34*, 124 – 130.

24. Wacker, R., Schroder, H., and Niemeyer, C. M. (2004) Performance of antibody microarrays fabricated by either DNA-directed immobilization, direct spotting, or streptavaidin-biotin attachment: a comparative study. *Anal. Biochem. 330*, 281 – 287.

25. Peluso, P., Wilson, D. S., Do, D., Tran, H., Venkatasubbaiah, M., Quincy, D., Heidecker, B., Poindexter, K., Tolani, N., Phelan, M., Witte, K., Jung, L. S., Wagner, P., and Nock, S. (2003) Optimizing antibody immobilization strategies for the construction of protein microarrays. *Anal. Biochem. 312*, 113 – 124.

26. Reichel, A., Schaible, D., Al Furoukh, N., Cohen, M., Schreiber, G., and Piehler, J. (2007) Noncovalent, site-specific biotinylation of histidine-tagged proteins. *Anal. Chem. 79*, 8590 – 8600.

27. Du, W., Ma, X., and Schneider, E. M. (2008) A direct immunoassay assessment of streptavidin- and *N*-hydroxysuccinimide-modified biochips in validation of serological TNFα responses in hemophagocytic lymphohistiocytosis. *J. Biomol. Screening 13*, 515 – 526.

28. Ro, H.-S., Jung, S. O., Kho, B. H., Hong, H. P., Lee, J. S., Shin, Y.-B., Kim, M. G., and Chung, B. H. (2005) Surface plasmon resonance imaging-based protein array chip system for monitoring a hexahistidine-tagge protein during expression an purification. *Appl. Environ. Microbiol. 71*, 1089 – 1092.

29. Kato, K., Sato, H., Kim, H. G., and Chung, B. H. (2005) Immbolization of histidine-tagged recombinant proteins onto micropatterned surfaces for cell-based functional assays. *Langmuir 21*, 7071 – 7075.

30. Wingren, C., Steinhauer, C., Ingvarsson, J., Persson, E., Larsson, K., and Borrebaeck, C. A. K. (2005) Microarrays based on activity-tagged single-chain Fv antibodies: Sensitive detection of analyte in complex proteomes. *Proteomics 5*, 1281 – 1291.

31. Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bildingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R. A., Gerstein, M., and Snyder, M. (2001) Global analysis of protein activities using proteome chips. *Science 293*, 2101 – 2105.

32. Khan, F., He, M., and Taussig, M. J. (2006) Double-hexahistidine tag with high-activity binding for protein immobilization, purification, and detection on Ni-Nitriloacetic acid surfaces. *Anal. Chem. 78*, 3072 – 3079.

33. Damrongchain, D., Yun, K., Kobatake, E., and Aizawa, M. (1997) Self-assembling of glutathione *S*-transferase:calmodulin fusion protein on chemically modified gold surface. *J. Biotech. 55*, 125 – 133.

34. Jung, J.-W., Jung, S.-H., Kim, H.-S., Yuk, J. S., Park, J.-B., Kim, Y.-M., Han, J.-A., Kim, P.-H., and Ha, K.-S. (2006) High-throughput analysis of GST-fusion protein expression and activity-dependent protein interactions on GST-fusion protein arrays with a spectral surface plasmon resonance biosensor. *Proteomics 6*, 1110 – 1120.

35. Kawahashi, Y., Do, N., Takashima, H., Tsuda, C., Oishi, Y., Oyama, R., Yonezawa, M., Miyamoto-Soto, E., and Yanagawa, H. (2003) *In vitro* protein microarrays for detecting protein-protein interactions: Application of a new method for fluorescence labeling of proteins. *Proteomics 3*, 1236 – 1243.

36. Yeo, W.-S., Min, D.-H., Hseih, R. W., Greene, G. L., and Mrksich, M. (2005) Label-free detection of protein-protein interactions on biochips. *Angew. Chem. Intl. Ed. Engl. 44*, 5480 – 5483.

37. Wu, S.-C., and Wong, S.-L. (2005) Engineering soluble monomeric streptavidin with reversible binding capability. *J. Biol. Chem. 280*, 23225 – 23231.

38. Lesaicherre, M.-L., Lue, R. Y. P., Chen, G. Y. J., Zhu, Q., and Yao, S. Q. (2002) Intein-mediated biotinylation of proteins and its application in a protein microarray. *J. Amer. Chem. Soc. 124*, 8768 – 8769.

39. Fischer, M., Leech, A. P., and Hubbard, R. E. (2011) Comparative assessment of different histidine-tags for immobilization of protein onto surface plasmon resonance sesnsorchips. *Anal. Chem. 83*, 1800 – 1807.

40. Valiokas, R., Klenkar, G., Tinazli, A., Tampe, R., Liedberg, B., and Piehler, J. (2006) Differential protein assembly on micropatterned surfaces with tailored molecular and surface multivalency. *ChemBiochem 7*, 1325 – 1329.

41. Harper, S., and Speicher, D. W. (2008) Expression and purification of GST fusion proteins. *Curr. Protocols Prot. Sci. Unit 6.6*, 6.6.1 – 6.6.26.

42. Ha, T. H., Jung, S. O., Lee, J. M., Lee, K. Y., Lee, Y., Park, J. S., and Chung, B. H. (2007) Oriented immobilization of antibodies with GST-fused multiple Fc-specific B-domains on a gold surface. *Anal. Chem. 79*, 546 – 556.

43. Hsu, K.-L., Gildersleeve, J. C., and Mahal, L. K. (2008) A simple strategy for the creation of a recombinant lectin microarray. *Mol. BioSys. 4*, 654 – 662.

44. Sudakevitz, D., Kostlanova, N., Blatman-Jan, G., Mitchell, E. P., Lerrer, B., Wimmerova, M., Katcoff, D. J., Imberty, A., and Gilboa-Garber, N. (2004) A new *Ralstonia*

solanacearum high-activity mannose binding lectin RS-IIL structurally resembling *Pseudomonas aeruginosa*. *Mol. Microbiol. 52*, 691 – 700.

45. Heron, B. T., Sateriale, A., Teixeira, J. E., and Huston, C. D. (2011) Evidence for a novel *Entamoeba histolytica* lectin activity that recognizes carbohydrates present on ovalbumin. *Int. J. Parasitol. 41*, 137 – 144.

46. Merckel, M. C., Tanskanen, J., Edelman, S., Westerlund-Wikstrom, B., Korhonen, T. K., Goldman, A. (2003) The structural basis of receptor-binding *Escherichia coli* associated with diarrhea and septicemia. *J. Mol. Biol. 331*, 897 – 905.

47. Sabin, C., Mitchell, E. P., Pokorna, M., Gautier, C., Utille, J.-P., Wimmerova, M. and Imberty, A. (2006) Binding of different monosaccharides by lectin PA-IIL from *Pseudomonas aeruginosa*: Thermodynamic data correlated with X-ray structures. *FEBS Letters 580*, 982 – 987.

48. Wu, A. M., Gong, Y.-P., Li, C.-C., and Gilboa-Garber, N. (2010) Duality of the carbohydrate-recognition system of *Pseudomonas aeruginosa*-II lectin (PA-IIL). *FEBS Lett. 584*, 2371 – 2375.

49. Blanchard, B., Nurisso, A., Hollville, E., Tetaud, C., Wiels, J., Pokorna, M., Wimmerova, M., Varrot, A., and Imberty, A. (2008) Structural basis of the preferential binding for globo-series glycosphingolipids displayed by *Pseudomona aeruginosa* lectin I. *J. Mol. Biol. 383*, 837 – 853.

50. Merckel, M. C., Tanskanen, J., Edelman, S., Westerlund-Wikstrom, B., Korhonen, T. K., Goldman, A. (2003) The structural basis of receptor-binding *Escherichia coli* associated with diarrhea and septicemia. *J. Mol. Biol. 331*, 897 – 905.

51. Huang, W., Yang, Q., Umekawa, M., Yamamoto, K., and Wang, L. X. (2010) Anthrobacter endo-β-*N*-acetylglucosaminidase shows transglycosylation activity on complex-type *N*-glycan oxazolines: one-pot conversion of ribonuclease B to sialylated ribonuclease C. *ChemBiochem 11*, 1350 – 1355.

52. Westerlund-Wikstrom, B. and Korhonen, T. K. (2005) Molecular structure of adhesin domains in *Escherichia coli* fimbriae. *Int. J. Med. Microbiol. 295*, 479 – 486.

53. Propheter, D. C., Hsu, K.-L., and Mahal, L. K. (2010) Fabrication of an oriented lectin microarray. *ChemBiochem 11*, 1203 – 1207.

54. Krishnamoorthy, L., and Mahal, L. K. (2009) Glycomic analysis: an array of technologies. *ACS Chem. Biol. 4*, 715 – 732.

Chapter 3: *In situ* orientation of GST-tagged lectins for glycomic analysis

3.1 Introduction

A major issue in protein microarray technology is the activity of the deposited protein. In some studies, groups demonstrated that as few as 5 – 20% of deposited protein samples show proper activity (*1* - *3*). As discussed in Chapter 2, the orientation of proteins *via* affinity tags can greatly enhance the binding sensitivity of a given microarray (*4*). In our original orientation work, we subdivided arrays on a microarray slide to generate dual-surface arrays. Although this methodology did allow both oriented and non-oriented proteins to be printed on the same array, it was technically challenging and impractical for use with complex samples such as membrane-bound glycans (*5*). We wondered whether we could simplify this two-step process into a single protein deposition and orientation step. We rationalized that supplementation of the print buffer with an excess of GSH in a slightly basic buffer might be able to outcompete protein immobilization through side chain lysines, allowing for a one pot *in situ* orientation scheme (Figure 3.1).

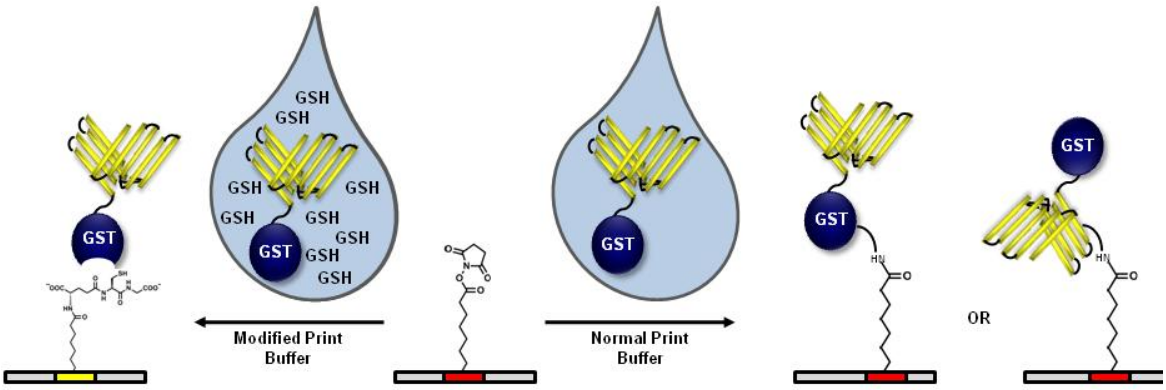

Figure 3.1 Schematic of the *in situ* orientation method. When printed under standard conditions, we obtain randomly deposited protein, thus affecting the activity of the immobilized protein. If we supplemented the print solution with an excess of GSH, we might be able to create a GSH-surface *in situ* which would organize GST-tagged proteins upon the GSH-scaffold (*16*).

We theorized the kinetics of the immobilization would favor the small molecule GSH over the protein, a large biomolecule. GSH is smaller, in excess, and contains a free terminal amine ($pK_a \sim 8$), thus it might outcompete the more bulky protein containing protonated lysine groups ($pK_a \sim 10$). If this were the case, when the protein solution is deposited onto the slide surface, the GSH would react first with the slide, creating a GSH-modified surface *in situ* which would organize the GST-tagged lectins upon it (Figure 3.1). To our knowledge, a microarray fabrication method for simultaneous slide derivatization and protein deposition has not been reported. Herein, we describe this new print methodology.

3.2 Results and Discussion

3.2.1 Optimization of print conditions for *in situ* orientation of GST-tagged lectins

For our initial experiments of *in situ* protein orientation, we chose to probe two lectins, GafD and RS-IIL, against Cy5-labeled chicken egg ovalbumin (OVA-Cy5), and just RS-IIL against Cy3-labeled RNase B (RNase B-Cy3). Based on our previous results in protein orientation, we tested both sodium borate (pH 8.5) and sodium bicarbonate (pH 9.3) buffers, as pH plays an important role in the amide coupling reaction between GSH and the NHS-activated slide. Once again, we analyzed overall lectin binding as the distribution of activities between the two lectins with respect to both buffer composition and amount of GSH additive (Figure 3.2).

We printed both GafD (20 μM) and RS-IIL (24 μM) in sodium borate and sodium bicarbonate buffers and in both buffers, a range of concentrations of GSH was tested (100, 50, 25, 10, and 0 mM). In Figure 3.2A, we plot the observed activity of GafD and RS-IIL against various concentrations of OVA-Cy5 with respect to each printing buffer containing all GSH concentrations. Again, just as in Chapter 2, we define activity as an increase in observed

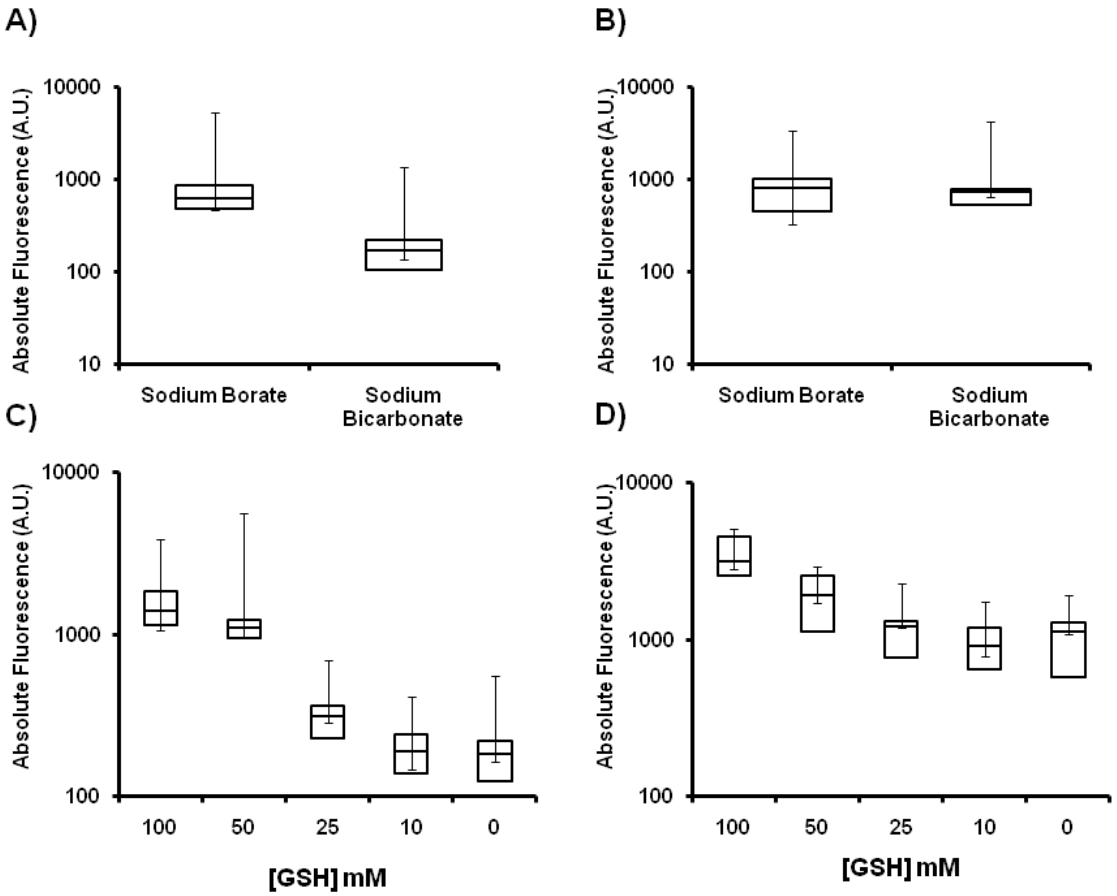

Figure 3.2 Optimization of GSH deposition for *in* situ orientation of GST-tagged lectins. Box and whisker plots representing the distribution of binding between two GST-tagged lectins GafD and RS-IIL, and two glycoproteins OVA-Cy5 and RNase B-Cy3. A) Distribution of binding of the two lectins against OVA-Cy5 in varying concentrations of GSH (100, 50, 25, 10, and 0 mM) with respect to the two print buffers sodium borate (pH 8.5) and sodium bicarbonate (pH 9.3). B) Distribution of binding of RS-IIL against RNase B-Cy3 with varying concentrations of GSH (100, 50, 25, 10, and 0 mM) with respect to the two print buffers sodium borate and sodium bicarbonate. C) Distribution of binding of the two lectins against OVA-Cy5 in both buffers with respect to GSH concentration. D) Distribution of binding of RS-IIL against RNase B-Cy3 in both buffers with respect to GSH concentration. From all the data collected, the optimized print buffer was found to be 100 mM GSH in sodium borate buffer (pH 8.5). Each box plot represents the distribution of binding in absolute fluorescence. The interquartile range (boxed) represents the distribution of 50% of the total binding signals and was used to compare between the different printing conditions (*16*).

fluorescence for a particular glycoprotein. In Figure 3.2B, we probed RS-IIL (24 μM) against RNase B-Cy3 using the same GSH concentrations as in A). In both conditions, the sodium borate buffer was found to be statistically better and different from the sodium bicarbonate values ($p <$

0.001). We initially believed that 50 mM GSH would be the optimal orientation concentration based on our previous work (*4*). When we analyzed the same data set to determine an optimal GSH concentration, the activities of GafD and RS-IIL against OVA-Cy5 at 100 and 50 mM GSH concentrations are very similar ($p = 0.7$) (Figure 3.2C). On the other hand, as shown in Figure 3.2D, the difference between the activities of RS-IIL against RNase B differ a significant amount between 100 and 50 mM GSH ($p < 0.001$). Upon closer examination of the data, we separated GafD and RS-IIL binding to OVA-Cy5 and noticed considerable differences. The difference between the 100 and 50 mM GSH concentrations for RS-IIL was statistically significant ($p = 0.025$), although for GafD they were not ($p = 0.96$). Thus we concluded that the optimal buffer for general GST-lectin activity was 50 mM sodium borate (pH 8.5) with 100 mM GSH. However, with the subtle differences in lectin binding activity between the two GSH concentrations for GafD, it may be possible that other proteins would require different buffer compositions. For the work presented in this chapter, our *in situ* orientation buffer (GSH-B, 50 mM sodium borate, pH 8.5, with 100 mM GSH, supplemented with 0.5 mg/mL BSA, final pH 4.2) was optimal for lectin activity compared to our standard print conditions (PB, 10 mM sodium phosphate, 15 mM sodium chloride, pH 7.4, supplemented with 0.5 mg/mL BSA).

This optimized buffer contains twice as much GSH as our buffer optimization for the production of our dual-surface arrays (*4*). We directly compared all three protein immobilization conditions to determine the effectiveness of both orientation methods compared to random deposition printing (Figure 3.3). RS-IIL (24 μM) was printed in GSH-B, PB, and then on a GSH-modified surface, and probed for activity against varying concentrations of OVA-Cy5

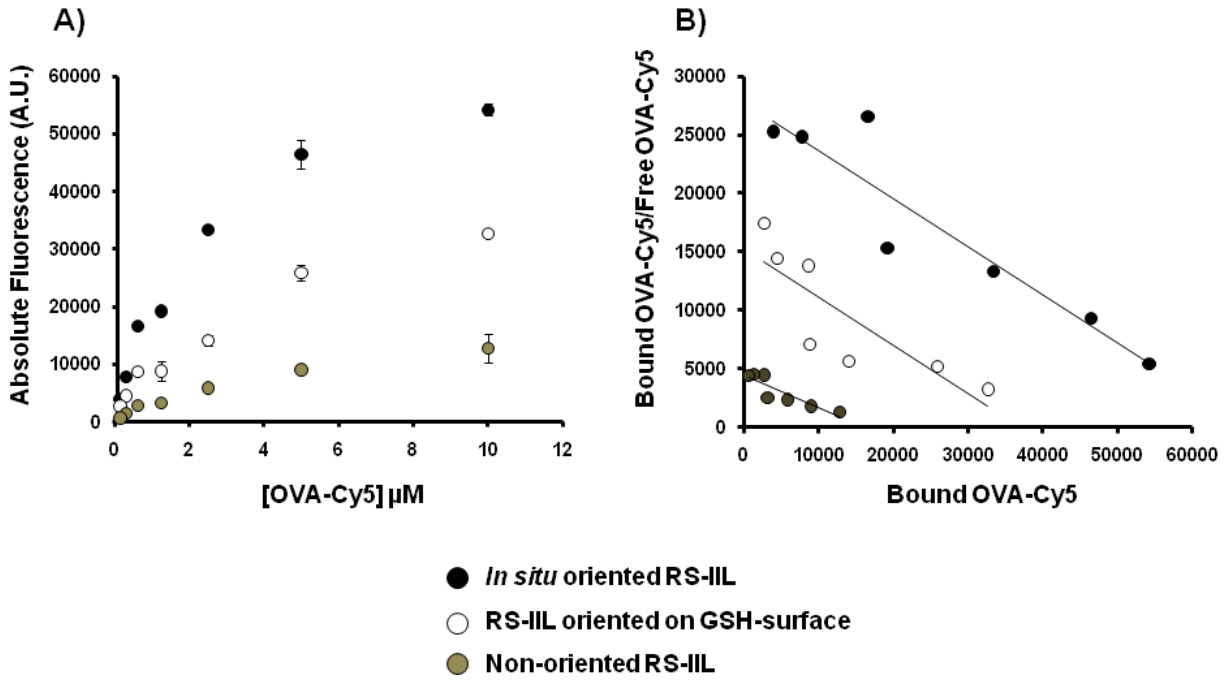

Figure 3.3 Comparison of the lectin orientation conditions between the *in situ* oriented and lectin oriented on GSH-surface. RS-IIL (24 µM) was printed in GSH-B (*in situ* oriented) and PB (non-oriented) on an NHS-activated surface. We also printed RS-IIL on a GSH-modified surface, and both slides were treated with varying concentrations of OVA-Cy5 (10, 5, 2.5, 1.25, 0.63, 0.31, and 0.16 µM). A) Activities of *in situ* oriented RS-IIL (closed circles), non-oriented RS-IIL (open circles), and RS-IIL oriented on a GSH-surface (tan circles) were plotted against the OVA-Cy5 concentration. B) Scatchard plots of RS-IIL binding to OVA-Cy5. Observed dissociation constants (K_d's) were similar in each deposition method. *In situ* oriented RS-IIL showed the greatest activity. Data is representative of triplicate arrays. Error bars indicate the standard deviation from the mean value (*16*).

(Figure 3.3). A direct comparison of the orientation methods revealed ~2-fold increase in activity in favor of *in situ* oriented RS-IIL over RS-IIL oriented on a pre-modified GSH-surface (Figure 3.3A). For further analysis, the data was plotted in the form of a Scatchard plot and the binding affinities were analyzed using linear regression to obtain apparent dissociation constants (K_d) (Figure 3.3B). The observed K_d's were similar to each other: (*In situ* oriented $K_d = 2.05 \pm 0.33$ µM, non-oriented $K_d = 3.59 \pm 0.39$ µM, and RS-IIL oriented on GSH-surface, $K_d = 2.63 \pm 0.47$ µM); and were in the typical range of binding activities for RS-IIL (*6*). The exact reason for the increase in activity between the two orientation methods is unknown, but the differences may arise from the different fabrication processes.

3.2.2 Activity of *in situ* oriented GST-tagged lectins

To demonstrate the effectiveness of *in situ* orientation, we printed GafD and RS-IIL in both GSH-B and PB and probed for activity against fluorescently-labeled glycoproteins, OVA-Cy5 and RNase B-Cy3, respectively.

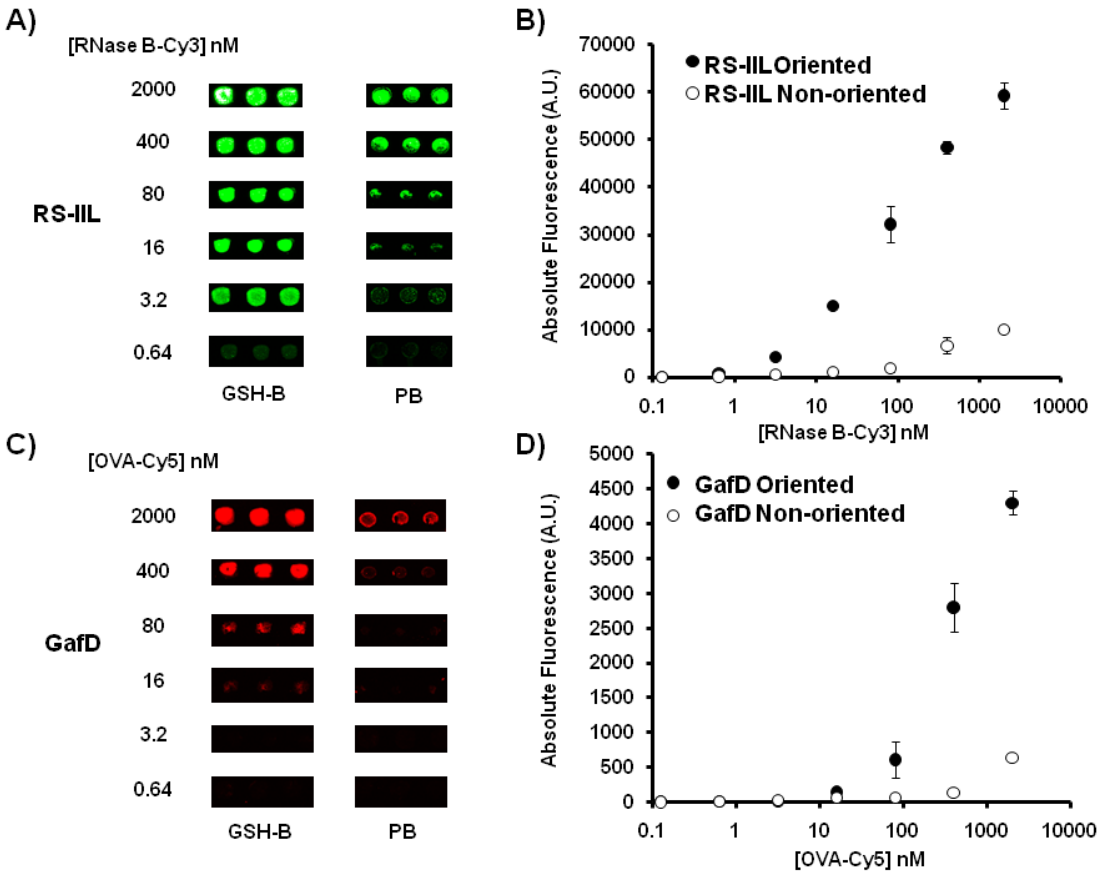

Figure 3.4 *In situ* orientation of recombinant lectins. A) Comparison of the activity of RS-IIL (24 μM) against varying concentrations of Cy3-labeled RNase B (RNase B-Cy3) printed in either GSH-B or PB. Data is representative of triplicate arrays. B) Graphical representation of data shown in A). C) Comparison of activity of GafD (10 μM) against varying concentrations of OVA-Cy5 printed in either GSH-B or PB. Data is representative of triplicate arrays. D) Graphical representation of data shown in C). Error bars indicate the standard deviation of the mean (*16*).

Both lectins displayed increased activity upon orientation. Moreover, improved spot morphology was observed for both lectins (Figure 3.4A and C). Under our standard print conditions, RS-IIL, the mannose-binding lectin, could detect RNase B-Cy3, a glycoprotein with a single high-mannose epitope (*7*), down to 16 nM. However, when oriented, we could detect the same glycoprotein down to the high picomolar range (640 pM), an increase of 25-fold in detection limits (Figure 3.4A and B). When GafD, the β-GlcNAc-binding lectin, was oriented *in situ*, an 9-fold increase in activity was observed against OVA-Cy5 versus the non-oriented lectin (Figure 3.4C and D). We also probed *in situ* oriented RS-IIL against OVA-Cy5 and observed a similar increase in activity against the mannose-containing glycoprotein. At the highest concentration of OVA-Cy5 tested (2 μM), we observed an ~8-fold increase in activity of *in situ* oriented RS-IIL over non-oriented RS-IIL (Figure 3.5). We also tested the effects of *in situ* orientation of PA-IIL, the fucose- and high-mannose binding lectin, against RNase B-Cy3. Consistent with our previous work (*4*), orientation increases glycan-binding activity is increased ~2-fold over several concentrations (Figure 3.6).

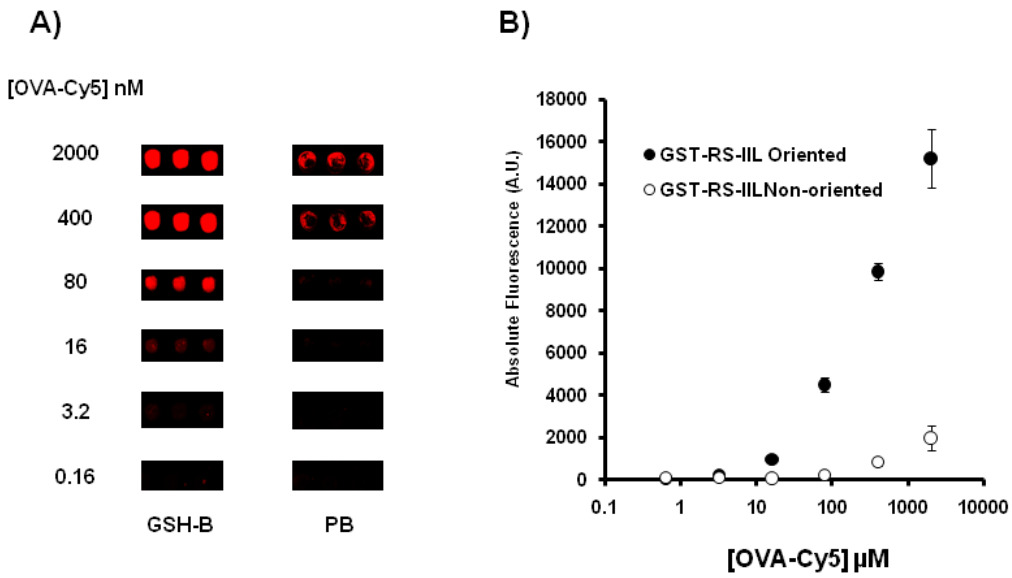

Figure 3.5 *In situ* orientation and activity of RS-IIL. A) RS-IIL (24 µM) was printed in either GSH-B (*in situ* oriented) or PB (non-oriented) and hybridized with Cy5-labeled ovalbumin (OVA-Cy5). B) Graphical representation indicates that the largest difference in activity is ~8-fold in favor of the oriented RS-IIL. The enhanced spot morphology of the oriented lectin is also clearly visible. Error bars indicate the standard deviation from the mean value. Data is representative of triplicate arrays (*16*).

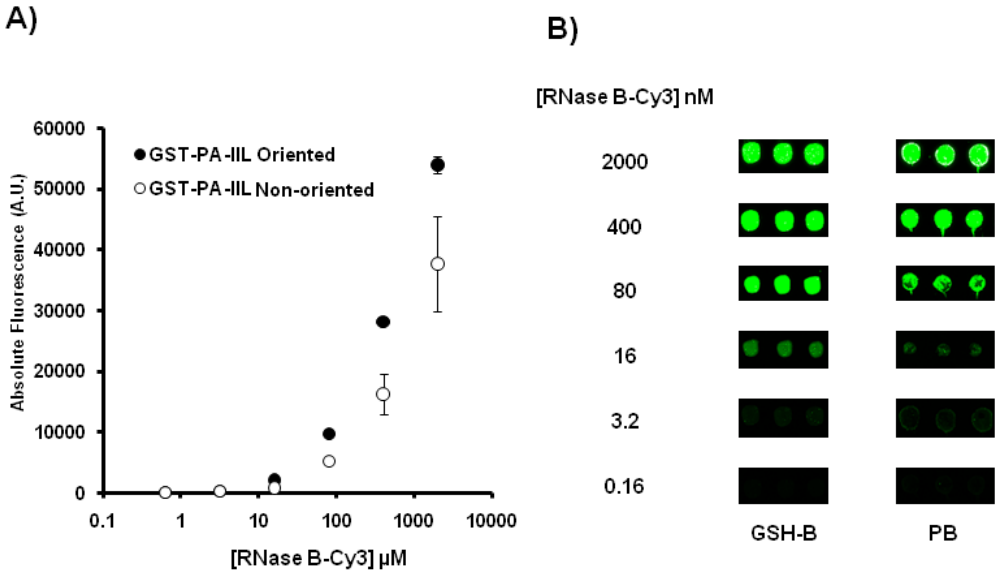

Figure 3.6 *In situ* orientation of PA-IIL. A) PA-IIL (24 µM) printed in either GSH-B (*in situ* oriented) or PB (non-oriented) and hybridized with varying concentrations of Cy3-labeled RNase B (RNase B-Cy3). Graphical representation of the data in B) indicates that binding of RNase B by PA-IIL shows modest improvement in activity (~1.5 – 2x) when oriented. Error bars indicate the standard deviation from the mean value. (b) Array data showing PA-IIL binding to RNase B. Data is representative of triplicate arrays (*16*).

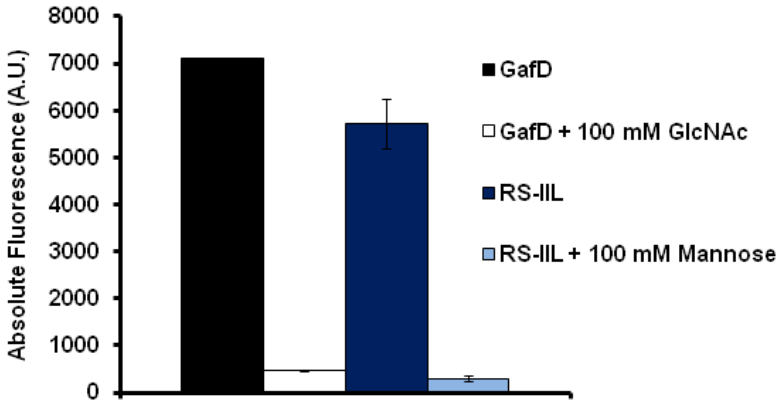

Figure 3.7 Competition of carbohydrate-binding of *in situ* oriented GafD and RS-IIL against OVA-Cy5 (10 μM). GafD (20 μM, black bar) displayed significant binding, yet when inhibited with 100 mM *N*-acetylglucosamine (GlcNAc, white bar), glycan-binding activity is diminished to background levels of binding. RS-IIL (24 μM, dark blue bar) showed significant binding, but when inhibited with 100 mM mannose (light blue bar), glycan-binding activity was decreased to background levels of binding.

In our original orientation work, we printed the same lectin-containing buffer on two different surfaces (*4*). In this work, we are dramatically altering the print buffer, therefore we wanted to determine that the lectin binding activity is not being affected by the excess GSH. In our previous work, we observed significant levels of background fluorescence on the GSH-modified surface. We wanted to know whether the signals observed after hybridization were carbohydrate-specific, and not due to the effects of a GSH monolayer. To test this, we printed GafD and RS-IIL in GSH-B and in PB, and then probed for activity against OVA-Cy5 in the absence or presence of inhibiting monosaccharide (Figure 3.7). Since GafD is a β-GlcNAc-binding lectin, we inhibited the carbohydrate interaction with 100 mM GlcNAc, and with RS-IIL, we inhibited the interaction with 100 mM mannose. The inhibition of GafD resulted in the loss of 93% of the fluorescence signal and the inhibition of RS-IIL resulted in a 95% decrease of fluorescence (Figure 3.7).

3.2.3 Testing the nature of the *in situ* orientation technique

To obtain a better understanding of both the limits and the mechanisms of *in situ* orientation, we performed several non-carbohydrate-based experiments. First, we fluorescently labeled BSA, a common additive in protein microarrays, with a maleimide-Alexa Fluor 647 conjugate and then printed in sodium borate buffer (3.5 μM Alexa Fluor 647 BSA in 50 mM sodium borate buffer, pH 8.5) with varying concentrations of GSH (Figure 3.8). To probe for deposition of GSH, we took advantage of the cysteine residue of the peptide and incubated the array with maleimide-PEG_2-biotin (250 μM), followed by incubation with Cy3-labeled streptavidin (streptavidin-Cy3, 50μg/mL). In the Cy3 channel, we observed deposition of the peptide at the lowest concentration tested (1 mM GSH, 180 A.U. (fluorescence) versus 0 mM GSH, 100 A.U.). Most notably, GSH deposition increases with GSH concentration, effectively out-competing the labeled BSA at higher concentrations (Figure 3.8). Given that our lectins are typically printed at ~20 μM, at the highest concentration, containing ~7 μM BSA and that both proteins contain ~30 lysines per protein, then we estimate that each lectin solution contains an ~800 μM concentration of lysine residues. At the 100 mM GSH concentration, we anticipate a molar excess of over 100-fold of GSH to lysine residues in our *in situ* orientation buffer. We cannot say with certainty the density of NHS-activated esters per spot volume, but the original slide developers concluded that 50 mM of a small amine was sufficient at deactivating the NHS-esters (*8*). In fact, the manual accompanying the Nexterion H slides suggests that, after protein printing, the slide should be blocked with 50 mM ethanolamine in sodium borate buffer (pH 8.5) (*9*). Therefore, we believe that at the 100 mM GSH concentration we can sufficiently react with all NHS-activated esters to form a GSH-modified surface *in situ*, thereby orienting GST-tagged proteins and repelling non-GST-tagged proteins.

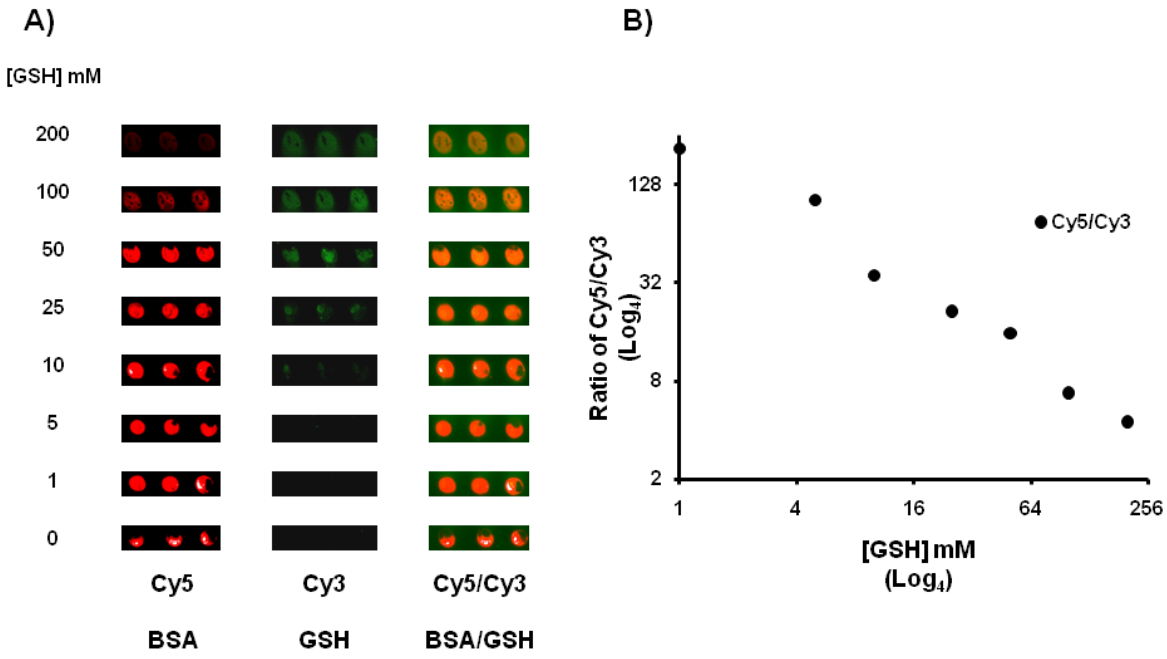

Figure 3.8 Competition of GSH and BSA for NHS-activated surface. Printing of Alexa Fluor 647-conjugated BSA (BSA-AF, 3.8 μM) in sodium borate buffer (pH 8.5) with varying amounts of GSH. A) Array data of BSA-AF printed in buffer containing GSH. The array was probed 250 μM maleimide-PEG_2-biotin, followed by 50 μg/mL of streptavidin-Cy3. Deposition of BSA-AF is observed in the Cy5 channel, and GSH deposition levels are observed in the Cy3 channel. An overlay of the spots is shown as a ratio (Cy5/Cy3). B) Graphical representation of data shown in A). Data is shown on a $\log_4$ scale to aid in visualization (*16*).

Given the increased activity of our GST-tagged lectins, we needed to determine whether the increased activity is simply due to increased levels of protein deposition. We printed Cy5-labeled GafD (GafD-Cy5) in both GSH-B and PB, and examined protein deposition and orientation by direct observation and indirect detection using the PE-conjugated α-S·tag antibody (α-S·tag-PE) (Figure 3.9). In our previous orientation method, we observed a ~40% decrease in deposition upon immobilization on the GSH-modified surface (*4*). When we printed GafD-Cy5 under these new set of conditions, we observed a 30% decrease in deposition of GafD-Cy5 when *in situ* oriented (Figures 3.9A and B). We also observed an increase in the indirect detection with α-S·tag-PE when GafD was oriented *in situ* (Figure 3.9C and D). These results are consistent with our previous findings on protein orientation as discussed in Chapter 2.

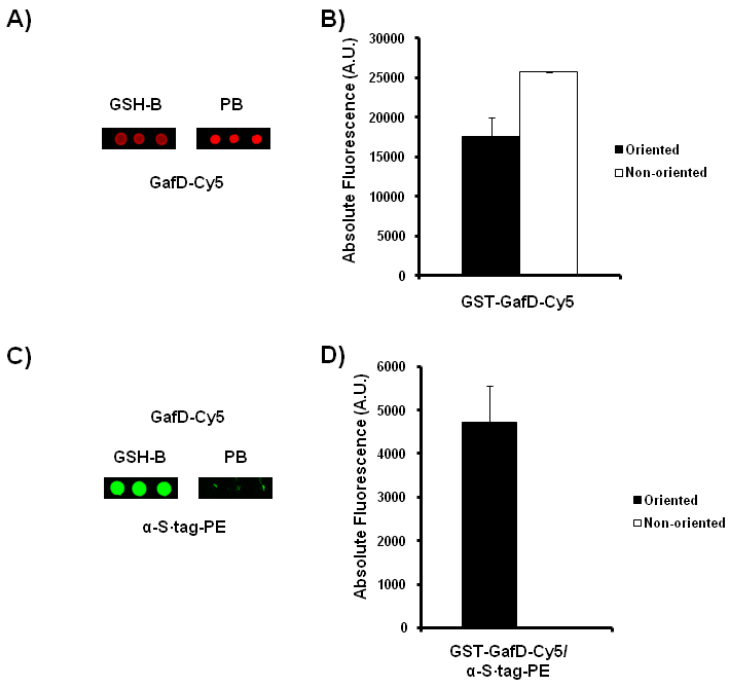

Figure 3.9 Comparison of the deposition and orientation of GafD-Cy5. A) GafD-Cy5 (10 µM) was printed in either GSH-B or PB. Array data is shown and representative of three separate arrays. B) Graphical representation of data shown in A). C) Indirect detection of GafD-Cy5 printed in GSH-B or PB and detected with α-S·tag-PE (1 µg/mL). Array data is shown and representative of three separate arrays. D) Graphical representation of data shown in C). The arrays spots shown are the same printed samples, just observed in two different emission channels. Error bars indicate the standard deviation from the mean value, and data is representative of triplicate arrays.

To quantify these results in terms of non-labeled GafD and the corresponding binding activity, we made a few assumptions in our calculation. First, all of the GafD immobilized and oriented in GSH-B binds to the glycoprotein. Second, the relative differences in activity of oriented versus non-oriented protein can be estimated from the differences in fluorescence signal from binding to OVA-Cy5. And third, the concentration of non-labeled and oriented GafD is also 30% lower than non-oriented GafD. Taking in these assumptions, we discover that only 8% of the non-oriented protein binds to the same glycoprotein (Calculation: (relative deposition/relative activity)*100%).

To demonstrate the necessity of the GST-domain for *in situ* orientation, we printed Cy5-labeled thrombin-cleaved GafD (tc-GafD-Cy5) (*4*) in both GSH-B and PB (Figure 3.10). When observing direct deposition in the Cy5 channel, we noticed a ~73% reduction of deposited tc-GafD-Cy5 when the protein was printed in GSH-B versus PB (Figures 3.10A and B). When we probed for the S·tag domain using α-S·tag-PE, we also observed a concomitant ~73% reduction in indirect detection of the deposited protein (Figures 3.10C and D). The equal reduction in both deposition and detection of an activity tag highlights that the necessity of a GST-tag for the *in situ* orientation.

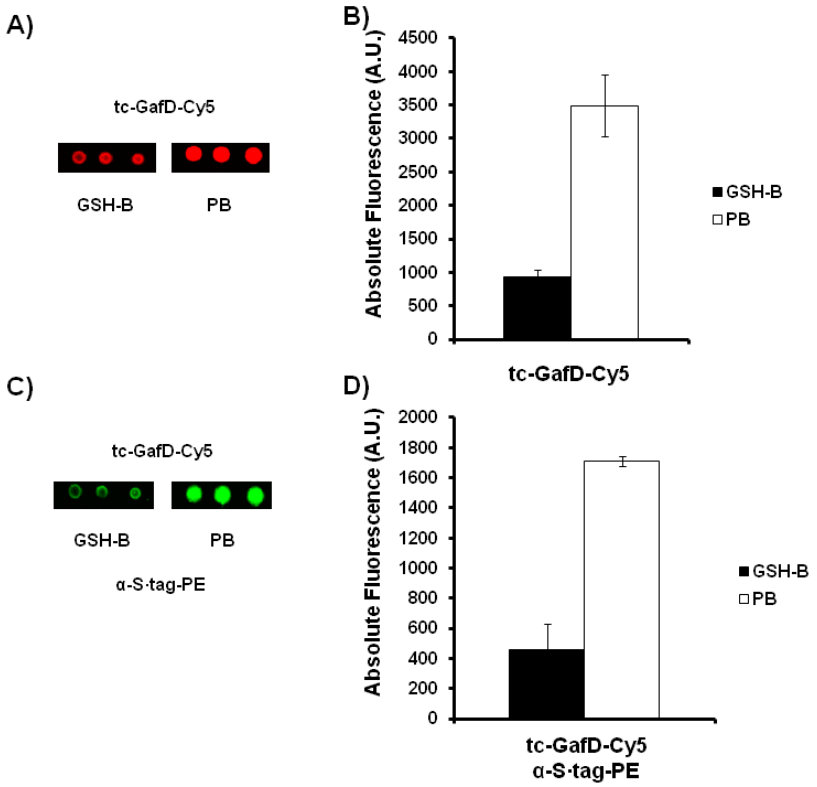

Figure 3.10 Comparison of the deposition and orientation of tc-GafD-Cy5. A) tc-GafD-Cy5 (10 µM) was printed in either GSH-B or PB. Array data is shown and representative of three separate arrays. B) Graphical representation of data shown in A). C) Indirect detection of tc-GafD-Cy5 printed in GSH-B or PB and detected with α-S·tag-PE (1 µg/mL). Array data is shown and representative of three separate arrays. D) Graphical representation of data shown in C). The arrays spots shown are the same printed samples, just observed in two different emission channels. Error bars indicate the standard deviation from the mean value, and data is representative of triplicate arrays.

The GST-GSH interaction is known to be a labile interaction in the presence of excess GSH, so we expected a facile elution of oriented GafD-Cy5 from the slide. However, when we attempted to remove deposited GafD-Cy5 using an excess of 1 M GSH, we observed moderate deposition prior to treatment with 1 M GSH, and saturating fluorescence after GSH treatment. Unfortunately, the Cy5 dye is very sensitive to reducing conditions (*10*), so the results were not interpretable. We also analyzed this phenomenon is solution, and observed an increase of fluorescence with respect to increasing GSH concentration. As an alternative, we incubated the array under denaturing conditions of 6 M urea for 1 hr. We would expect that the GST-GSH interaction would be disrupted upon denaturation of the GST domain, and a comparison of a before and after treatment of urea showed a 50% loss in GafD-Cy5 deposition (Figure 3.11). In contrast, GafD-Cy5 that was randomly immobilized onto the NHS-activated surface displayed no decrease in fluorescence upon treatment with urea (Figure 3.11). Our observation indicates that at least 50% of our *in* situ oriented protein is non-covalently bound to our GSH surface. However, half of all protein was still present after the denaturing conditions.

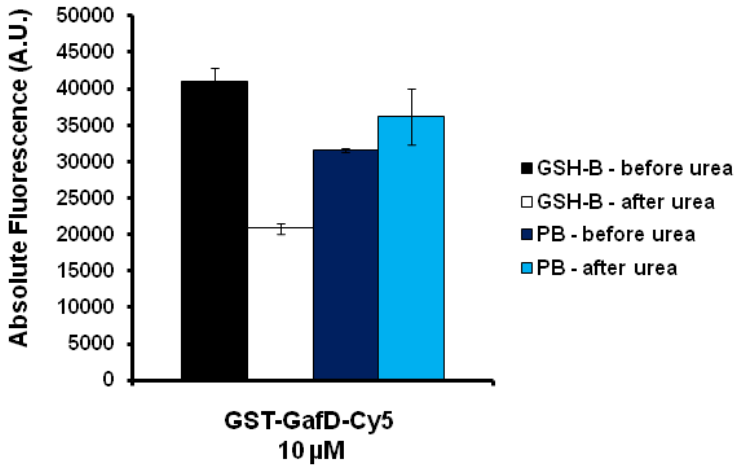

Figure 3.11 The effects of denaturation of *in situ* oriented and randomly immobilized GafD-Cy5 . GafD-Cy5 (10 µM) was printed in both GSH-B and PB, scanned (before urea), and then treated with 6 M urea, and re-scanned at the same conditions (after urea). Error bars indicate the standard deviation from the mean value. Data is representative of duplicate arrays (*16*).

There may be several possible reasons for this result. First, it is plausible that the GSH binding domain of GST is buried within the Nexterion H surface, which may protect it from denaturation by urea. When in a dry state, the NHS-PEG surface of the slide is ~10 nm thick. When hydrated, the surface expands to between 50 – 100 nm, thus allowing adsorption of the sample into the slide (*8*). Considering that proteins, in general, are well within a ~100 nm diameter, the GSH binding domain could be buried within the Nexterion H slide, protecting the protein and also making the binding interaction difficult to compete out. Another possible reason may be that the remaining GafD is covalently bound to the surface. In this instance, the GSH may still initially form a GSH monolayer, yet some NHS-esters still remain, thus leaving those remaining amine-reactive functional groups available for lysine conjugation. In both cases, we expect an initial formation of a GSH monolayer.

When testing the pH of the final solution of GSH-B, we observed some intriguing results. First, as expected, the 50 mM sodium borate buffer alone and with additives Tween 20 (0.001%) or BSA (0.5 mg/mL) had very similar pH values (8.58, 8.57, and 8.54, respectively). However, with the addition of 100 mM GSH, we lower the pH down to 4.22, creating a more acidic environment than previously expected. To determine whether the final pH of the buffer is important for *in situ* orientation, we printed GafD (10 μM) in both the standard buffer, GSH-B pH 4.2, and in a buffered solution, GSH-B pH 8.5, against Cy3-labeled GlcNAc-BSA (GlcNAc-BSA-Cy3, Figure 3.12). In fact, when observing the activity over multiple concentrations of GafD (40, 20, 10, 5, and 2.5 μM) and multiple concentrations of GlcNAc-BSA-Cy3 (2000, 400, 80, 16, 3.2, 0.64, 0.128, and 0.256 nM), the activity of *in situ* oriented GSH-B pH 4.2 is statistically different than GSH-B pH 8.5 ($p < 0.001$). Although the amines in solution are most likely protonated at this low pH, the mechanism of immobilization and orientation should remain

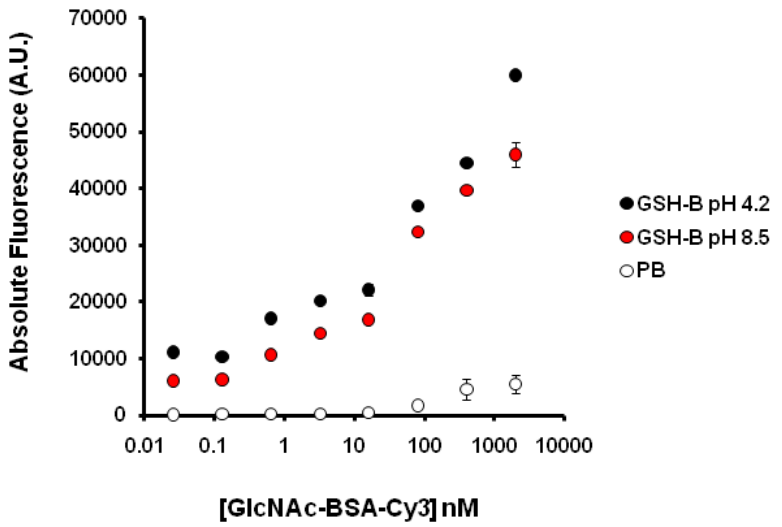

Figure 3.12 Comparison of activity between GafD printed in two different GSH-Bs. GafD was printed in the original print buffer GSH-B has a pH of 4.2 (closed circles), a modified GSH-B was buffered to pH 8.5 (red circles), and our standard PB was maintained at pH 7.4 (open circles).

the same, since the relative pK_a's are still the same, meaning the GSH molecule will still outcompete the the lysines on GST-tagged GafD. Also, since we observed enhanced spot morphology with our *in situ* orientation method, ie no "coffee cup effect" (*11*), the excess GSH competes with the protein for interactions with the air-water interface resulting in more uniform protein deposition.

3.3 Conclusions

Herein we present a new method for the orientation of recombinant, GST-fusion proteins *in situ* on a solid support. By taking advantage of the difference in reaction kinetics between our proteins and GSH for amide bond formation with an NHS-ester, we have created a one-pot method for the fabrication of a localized GSH scaffold and the organization of GST-tagged proteins upon it. After some optimization, we found that we could supplement our print buffer with GSH, which could modify the NHS-activated slide surface *in situ* that could then orient our GST-tagged lectins. We observed increased activity to a similar extent of our lectins printed on

our dual-surface array, and we showed that we are indeed orienting our lectins *in situ*. Our method maintains the native slide chemistry of the NHS-activated slide, allowing us to maintain the diversity of lectins afforded to us from the naturally derived proteins, crucial to the utility of our arrays in glycomic analysis. In this study, we only analyzed orientation by using reduced glutathione, however, since GST can still bind to oxidized GSH (GSSG) (*12*), it may be possible to use this ligand as well. Although we apply our orientation strategy to GST-tagged recombinant lectins, there is little to suggest that this method cannot be widely applied to other GST-tagged proteins or even, with some modification, to other protein tag systems. Indeed, it may be possible to orient the other two most widely used activity tags, biotin- and His_6-tags. For biotinylated substrates, we might be able to synthesize a GST-streptavidin fusion protein. When printed in GSH-B and with a biotin-tagged protein, the GSH could orient the GST-streptavidin, which would then orient a biotinylated protein. As for the His_6-tag, it may be possible to orient a His_6-tagged protein using a Ni^{2+}-NTA mixture. Indeed, current efforts in the group indicate that we could orient a multi-His-tagged (*13*) lectin with a Ni^{2+}-NTA-containing buffer. Lectin microarray technology is shifting toward the implementation of more recombinant lectins. Aside from our work, another group has attempted to integrate recombinant lectins into their format (*14*). Integrating new recombinant lectins with common fusion tags requires the corresponding development of newer orientation methods.

3.4 Materials and methods

3.4.1 General microarray fabrication

Recombinant lectins were expressed and purified as described previously (*15*) GafD-Cy5 was expressed, purified, and labeled as previously described (*4*) Unless otherwise noted, all microarrays were printed *via* the following protocol: Lectins were diluted in either GSH-B (50

mM sodium borate buffer, pH 8.5, containing 100 mM GSH) or PB (phosphate buffered saline (PBS), 10 mM sodium phosphate, 15 mM sodium chloride). For lectin printing, both print buffers contain 1 mM $CaCl_2$, 1 mM $MgCl_2$, and 1 mM of appropriate monosaccharide (Appendix A Table 2), and 0.5 mg/mL BSA. Prepared samples were loaded into a 384-well microplate (Whatman, Piscataway, NJ), and loaded into the SpotBot2 Personal Microarrayer (ArrayIt, Sunnyvale, CA). Printing programs were created with the MMF Spocle Program. Samples were printed onto Nexterion H slides (Schott North America, Elmsford, NY) with an SMP3 pin (ArrayIt, #SMP3). During the print, the slides were kept at 8 °C with internal humidity maintained at ~50% throughout the print process. After printing, the slides were allowed to warm to room temperature for 2 hr, while maintaining humidity at ~50%. After 2 hr, the slides were then placed in a coplin jar and blocked with 50 mM ethanolamine in 50 mM sodium borate buffer (pH 8.5) for 1 hr, at room temperature with mild shaking. After one hour, the slides were washed with PBS with 0.05% Tween (0.05% PBS-T, 3 x 3 min) and once with PBS. The slides were dried using a slide spinner (Labnet Intl., Edison, NJ), and then fastened in a 24-well hybridization chamber (ArrayIt). Fluorescently labeled samples were diluted into 0.005% PBS-T, and 100 µL were added to each subarray and samples were incubated for 2 hr at room temperature with gentle shaking. For visualization with α-S·tag-PE antibody, the buffer was changed to 0.005% PBS-T with 1% BSA. After 2 hr, samples were aspirated and washed with 0.005% PBS-T (0.005% Tween 20 in PBS, 3 x 3 min) and once with PBS. The slides were dried as before and loaded into the Genepix 4100A slide scanner (Molecular Devices, Union City, CA). Data was extracted with GenePixPro 5.0 (Molecular Devices) and analyzed and graphed using Microsoft Office Excel 2007.

For the buffer optimization, we printed GafD (20 μM) and RS-IIL (24 μM) versus two OVA-Cy5 and RNase B-Cy3. The lectins were printed in either 50 mM sodium borate buffer (pH 8.5) or 100 mM sodium bicarbonate (pH 9.3) supplemented with various concentrations of GSH (100, 50, 25, 10, and 0 mM). After immobilization and blocking, the slides arrays were incubated with varying amounts of OVA-Cy5 and RNase B-Cy3 (10, 5, 2.5, 1.25, 0.625, and 0.313 μM). The box and whisker plots were generated using Microsoft Excel 2007.

3.4.2 Labeling and printing of BSA-AF 647

Alexa Fluor® 647 C_2-maleimide (667 μM, dissolved in PBS, Invitrogen #A20347) was added to BSA (10 mg/mL, ~150 μM) in PBS and incubated at room temperature for 1 hr with gentle shaking. After 1 hr, the sample was dialyzed against PBS for 12 hr, and the final concentration was determined by DC Assay (Bio Rad #500-0112). BSA-AF (3.8 μM) was dissolved in 50 mM sodium borate buffer (pH 8.5) with varying amounts of GSH (200, 100, 50, 10, 5, 1, and 0 mM) and printed as described above. After blocking, washing, and drying the slide, 100 μL of 250 μM Maleimide-PEG_2-Biotin (Thermo Fisher Scientific, #21901), in 0.005% PBS-T was added to the array and incubated at room temperature for 1 hr. After 1 hr, the slide was washed with 0.005% PBS-T (3 x 3 min) and once with PBS, and dried. 100 μL of 50 μg/mL of Cy3-labeled streptavidin (Invitrogen, #43-4315) was then added to each well and incubated for 1 hr at room temperature. After 1 hr, slides were then washed, dried, and scanned as previously described.

3.4.3 Urea treatment of immobilized lectins

GST-GafD was printed in 50 mM sodium borate (pH 8.5) containing 100 mM GSH (GSH-B) on Nexterion H slides as described above. After printing, the slides were allowed to incubate for two hours, equilibrating to room temperature. The slides were then blocked with 50

mM ethanolamine in 50 mM sodium borate buffer (pH 8.5) for 1 hr, then washed with 0.05% PBS-T (3 x 3 min) and once with PBS. The slides were dried and fastened to a 24-well hybridization chamber (ArrayIt). Arrays were then either treated to 100 µL of 6 M urea in PBS or 0.005% PBS-T, and were incubated for 1 hr at room temp. After 1 hr, the slides were washed with 200 µL of 0.005% PBS-T (3 x 3 min), then were treated with either 10 µM Cy5-labeled ovalbumin (OVA-Cy5) or α-S·tag antibody PE-conjugated in 0.005% PBS-T or 0.005% PBS-T with 1% BSA, respectively. The slides were then washed with 0.005% PBS-T (3 x 3 min) and once with PBS. The slides were dried, scanned, and analyzed as described above.

3.5 References

1. Wingren, C., and Borrebaeck, C. A. K. (2004) High-throughput proteomics using antibody microarrays. *Expert. Rev. Proteomics 1*, 355 – 364.
2. Wingren, C., and Borrebaeck, C. A. K. (2006) Antibody microarrays: current status and key technological advances. *Omics 10*, 411 – 427.
3. Borrebaeck, C. A. K., and Wingren, C. (2009) Design of high-density antibody microarrays of disease proteomics: Key technological issues. *J. Proteomics 72*, 928 – 935.
4. Propheter, D. C., Hsu, K.-L., and Mahal, L. K. (2010) Fabrication of an oriented lectin microarray. *ChemBiochem 11*, 1203 – 1207.
5. Hsu, K.-L., Pilobello, K., Krishnamoorthy, L., and Mahal, L. K. (2011) Ratriometric lectin microarray analysis of the mammalian cell surface glycome. *Meth. Mol. Biol. 671*, 117 – 131.
6. Sudakevitz, D., Kostlanova, N., Blatman-Jan, G., Mitchell, E. P., Lerrer, B., Wimmerova, M., Katcoff, D. J., Imberty, A., and Gilboa-Garber, N. (2004) A new *Ralstonia solanacearum* high-activity mannose binding lectin RS-IIL structurally resembling *Pseudomonas aeruginosa*. *Mol. Microbiol. 52*, 691 – 700.
7. Liang, C. J., Yamashita, K., and Kobata, A. (1980) Structural study of the carbohydrate moiety of bovine pancreatic ribonuclease B. *J. Biochem. 88*, 51 – 58.
8. Harbers, G. M., Emoto, K., Greef, C., Metzger, S. W., Woodward, H. N., Mascali, J. J., Grainger, D. W., and Lochhead, M. J. (2007) Functionalized Poly(ethylene glycol)-based bioassay surface chemistry that facilitates bio-immobilization and inhibits nonspecific protein, bacterial, and mammalian cell adhesion. *Chem. Mater. 19*, 4405 – 4414.

9. Schott North America Inc. (2009) Protocol: Nexterion® slide H protein application. http://www.us.schott.com/nexterion.

10. Berlier, J. E., Rother, A., Buller, G., Bradford, J., Gray, D. R., Filanoski, B. J., Telford, W. G., Yue, S., Liu, J., Cheung, C.-Y., Chang, W., Hirsch, J. D., Beechem, J. M., Haugland, R. P., and Haugland, R. P. (2003) Quantitative comparison of long-wavelength Alexa Fluor dyes to Cy dyes: Fluorescence of the dyes and their bioconjugates. *J. Histochem. Cytochem. 51*, 1699 – 1712.

11. Deng, Y., Zhu, X.-Y., Kienlen, T., and Guo, A. (2006) Transport at the air/water interface is the reason for rings in protein microarrays. *J. Amer. Chem. Soc. 128*, 2768 – 2769.

12. Nishihara, T., Maeda, H., Okamoto, K.-I., Oshida, T., Mizoguchi, T., and Terada, T. (1991) Inactivation of human placenta Glutathione-*S*-Transferase by SH/SS exchange reaction with biological sulfides. *Biochim. Biophys. Res. Comm. 174*, 580 – 585.

13. Fischer, M., Leech, A. P., and Hubbard, R. E. (2011) Comparative assessment of different histidine-tags for immobilization of protein onto surface plasmon resonance sesnsorchips. *Anal. Chem. 83*, 1800 – 1807.

14. Tateno, H., Toyota, M., Saito, S., Onuma, Y., Ito, Y., Hiemori, K., Fukumura, M., Matsushima, A., Nakanishi, M., Ohnuma, K., Akutsu, H., Umezawa, A., Horimoto, K., Hirabayashi, J., Asashima, M. (2011) Glycome diagnosis of human induced pluripotent stem cells using lectin microarray. *J. Biol. Chem. 286*, 20345 – 20353.

15. Hsu, K.-L., Gildersleeve, J. C., and Mahal, L. K. (2008) A simple strategy for the creation of a recombinant lectin microarray. *Mol. BioSys. 4*, 654 – 662.

16. Propheter, D.C. and Mahal, L. K. (2011) Orientation of GST-tagged lectins *via in situ* surface modification to create an expanded elctin microarray for glycomic analysis. *Mol. BioSys. 7,* 2114 – 2117.

Chapter 4: *In situ* orientation of non-tagged IgG and IgM antibodies

4.1 Introduction

Despite the widespread use of lectin microarrays, the majority of protein microarrays produced today employ antibodies as the antigen-binding proteins (*1* - *5*). Antibodies, which typically bind in the nM to pM K_d range (*3*), are ideal supplements to the lectin microarray for glycomic profiling. Antibody microarrays are typically printed onto highly adsorptive surfaces, such as nitrocellulose slides, which can often give high background fluorescence (*6*). Previous work in the Mahal lab indicated that a small set of antibodies printed on the Nexterion H slides gave poor activity. This may be a result of using an inactive antibody or, since protein immobilization can result in aggregation or blocked antigen-binding sites, greatly diminished activity (*3*). With the successful orientation of GST-tagged lectins *in situ*, we wondered whether we could apply this technique to the immobilization of antibody-binding proteins that would then orient the antibodies *in situ*.

The most prevalent and commercially-available carbohydrate-binding antibodies are IgG and IgM antibodies (Figure 4.1). The IgG class of antibodies is the most abundant in the body (~75%) (*7*), and is subdivided further into four subclasses (IgG_1, IgG_2, IgG_3, and IgG_4), which vary structurally in the composition of the Fc domains (*8*). With respect to the IgM antibodies, IgG antibodies are monomeric in nature, meaning they comprise two antigen-binding domains (F_{Ab}) tethered to a single constant domain (F_c) (*9*). In contrast, IgM antibodies are pentameric when secreted, comprising of five antibody monomers with a total of ten antigen-binding domains (*10*). The monomers are all connected through the F_c domains *via* disulfide bridges to a central peptide, called the J chain (*11*).

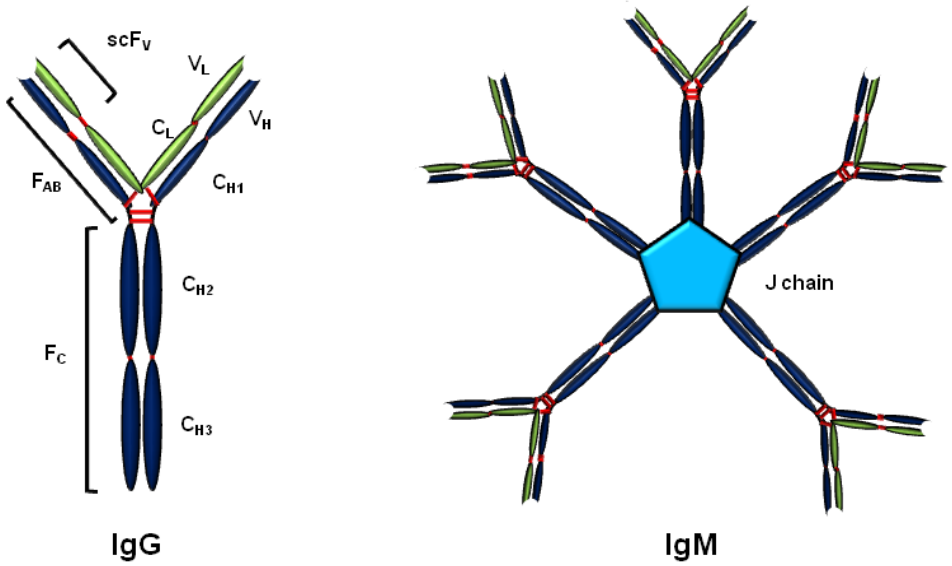

Figure 4.1 Representations of the IgG and IgM class of antibodies. The most abundant class of antibodies, IgG, are monomeric antibodies with two antigen binding domains located at the tip of the variable regions of heavy and light chains (V_H and V_L). Secreted IgM antibodies are pentameric antibodies which contain 10 antigen binding domains, connected at the end of the individual Fc domains by a J chain peptide.

Protein A (SpA) and protein G (SpG) are well-characterized antibody-binding proteins secreted by the bacterial families *Staphylococcus* and *Streptococcus*, respectively (*12,13*). Both antibody-binding proteins bind to the F_c domains, but SpG can also bind to the constant regions of the F_{ab} to a minimal degree (*14*). Groups that have reported on the orientation of IgG antibodies using SpG, simply assume that the F_c domain is bound and not the F_{ab} due to the differences in activity. The first two groups that attempted antibody orientation directly immobilized SpA to activated slide, such as epoxy- or aldehyde-modified surfaces (*15 - 17*). Wang et al. demonstrated the immobilization of SpA on a gold nanoparticle surface in a flow-reaction chamber to use in an antibody-antigen binding system (*15*). Matson et al. describe a method to orient antibody microarrays using a print-over-print method, which requires two separate print procedures (*16*). One widely-cited work on protein orientation regards the orientation of F_c-fused dectin-1 (*18*). Chen et al. tested three orientation methods, two of which

used SpG and boronate esters, to bind to the *N*-glycans present on the F_c domain of the antibodies. The authors end up not using SpG, but it highlights the need for a general method for antibody orientation in a microarray format (*18*). A major issue with these methods is that not one discusses any potential benefits of the immobilization technique of the SpA/G. It seems obvious that improved antibody orientation can be achieved with an oriented SpA/G. To address this issue, another group, whose work is discussed in the Chapter 2, synthesized a GSH-containing small molecule which they coated onto a gold surface. GST-tagged SpG derivatives were then printed onto the modified slide and binding activities to IgG antibodies were assessed (*19*). Although a good example for the need of oriented SpG, the synthesis of the small molecule is not practical. Furthermore, since entire slide chemistries are specialized to a specific protein, there is a loss of diversity in the potential of arrayed proteins, which severely limits the application of this method.

Since we could orient single GST-tagged protein *in situ*, we wondered whether we could print GST-tagged SpA or SpG with an IgG antibody in our GSH-containing buffer. Unfortunately, only about half of all commercially-available carbohydrate-binding antibodies are IgG, the rest are IgM antibodies. Given the pentavalent nature of IgM antibodies, we hypothesized that the F_c domains would be inaccessible for binding by the traditional antibody-binding proteins SpA and SpG. For this reason, we generated a GST-tagged version of Protein L (PpL), an antibody-binding protein derived from *Peptostreptococcus magnus*, which binds the constant region of κ light chain (*20*). Although available as a resin for the purification of antibodies (Pierce, Rockford, IL), PpL has never been used to orient IgM antibodies nor in protein microarray fabrications. When PpL and an IgM antibody are printed in GSH-B, the GSH

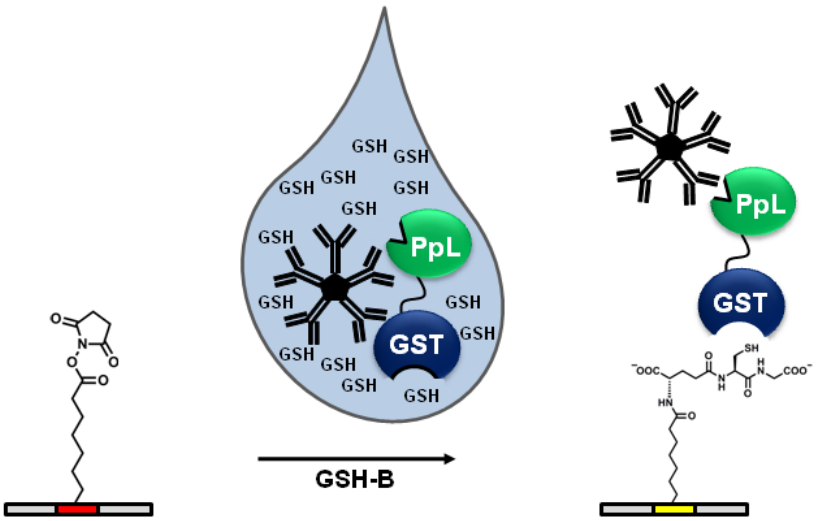

Figure 4.2 Schematic for the *in situ* orientation of IgM antibodies using GST-PpL. When the solution is printed, the GSH reacts first with the slide surface, forming the initial GSH-monolayer. Then the GST-tagged antibody binding protein GST-PpL is oriented onto the GSH-surface, which tethers the IgM protein in an oriented manner. IgG antibodies are printed in the same way except for the use of GST-tagged SpA and SpG.

forms a small monolayer which would orient the GST-PpL fusion protein, which in turn would orient the antibody *via* the κ light chains (Figure 4.2). For IgG antibodies, the orientation scheme is identical except that the GST-tagged versions of SpA and SpG orient the antibody *via* the F_c domain. We are currently exploring this method and the possible applications in generating a diverse library of CBPs for glycomic analysis.

4.2 Results and Discussion

4.2.1 *In situ* orientation of IgG antibodies

In order to orient IgG antibodies using the GSH-based orientation scheme, I cloned the antibody-binding fragments of Protein A and G from either commercially available genomic DNA (SpA) or from synthesized DNA (SpG, See Appendix B for gene), and ligated the fragments into the pET41 Ek/LIC kit (Novagen). After transformation and expression, the antibody-binding proteins were purified *via* the GST-domains (Figure 4.3).

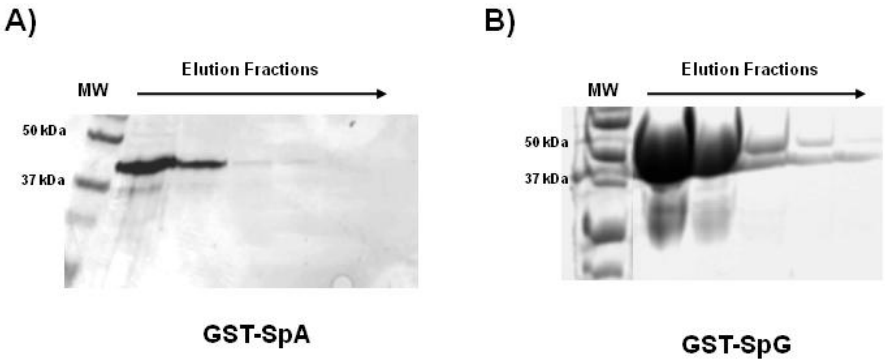

Figure 4.3 Purified GST-tagged antibody-binding proteins SpA, and SpG. Proteins were expressed in BL21 DE3 cells and purified *via* GSH-sepharose columns. A) 10% SDS-PAGE gel analysis of SpA, B) SpG, and Both GST-tagged SpA and SpG proteins are ~ 44 kD in size.

After purification of GST-tagged SpA and SpG, our initial work focused on optimizing the antibody to SpA/G ratio using the *in situ* orientation method. To do this, we developed a two-color orientation experiment with orthogonally labeled antibodies. SpA (125 μg/mL) would dissolved in GSH-B (see Chapter 3) and mixed with varying concentrations of Cy5-labeled goat α-mouse IgG. The immobilized antibody would be probed for activity against a Cy3-labeled mouse IgG antibody (Figure 4.4). To compare to the affect of orientation on the activity of deposited antibody, we printed the Cy5-labeled antibody in our standard print buffer (PB, see Chapter 3). With this method, we will be able to directly compare the amount of deposited antibody versus the resultant activity under *in situ* orientation conditions and print conditions that result in the random immobilization of the antibody.

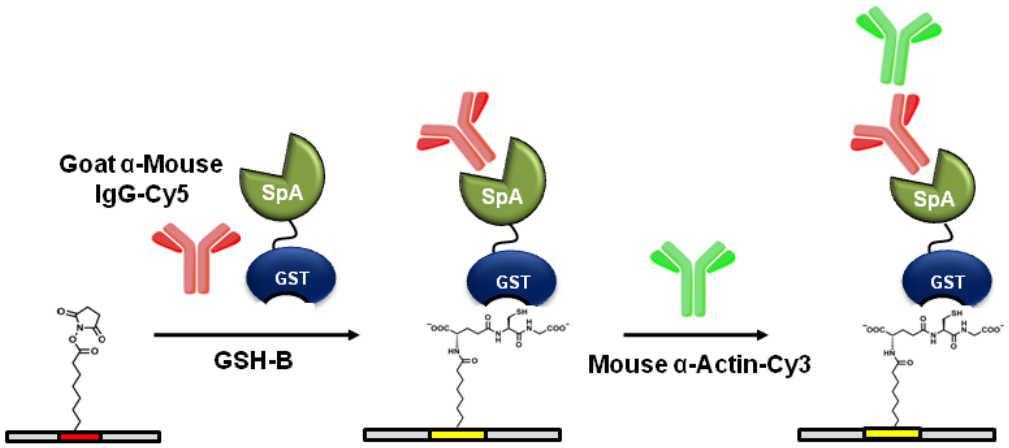

Figure 4.4 Experimental design of a dual color antibody orientation format. Cy5-labeled Goat α-mouse IgG antibody is printed with GST-SpA in GSH-B. Deposition of the protein solution results in the *in situ* derivitization of the NHS-activated surface. Upon this GSH-layer, the GST-tagged SpA is oriented, which then binds to the Fc domain of the IgG antibody. The oriented Cy5-labeled antibody can then more effectively bind to a Cy3-labeled mouse IgG.

To control for binding of the secondary antibody to the oriented GST-SpA, we analyzed the binding of various concentrations of Cy5-labeled goat α-mouse IgG versus a single concentration of GST-SpA (125 µg/mL). At the concentrations tested, we observed that at about a 1:1 molar ratio of antibody:GST-SpA, we observe a saturation of binding (500 µg/mL antibody : 125 µg/mL GST-SpA) (Figure 4.5). At this concentration, when directly compared to Cy5-labeled antibody printed in PB, the oriented antibody displayed ~13-fold decrease in overall antibody deposition (Figure 4.6A and B). To compare the activity of the oriented antibody, we probed the arrays with Cy3-labeled mouse IgG antibody (50 µg, α-actin). Given the substantial amount of randomly immobilized antibody compared to *in situ* oriented antibody, we were surprised to observe an increase of 7-fold more activity for the oriented antibody (Figure 4.6A and B).

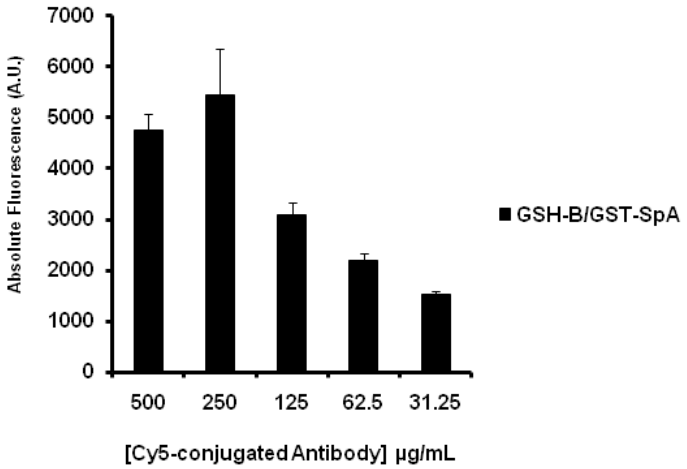

Figure 4.5 Determination of the saturation of the antibody-binding sites of SpA against varying concentrations of Cy5-labeled IgG. In this buffer composition (GSH-B/GST-SpA), at a 1:1 molar concentration of antibody and GST-SpA, we observe a saturation of binding (500 µg/mL antibody:125 µg/mL GST-SpA).

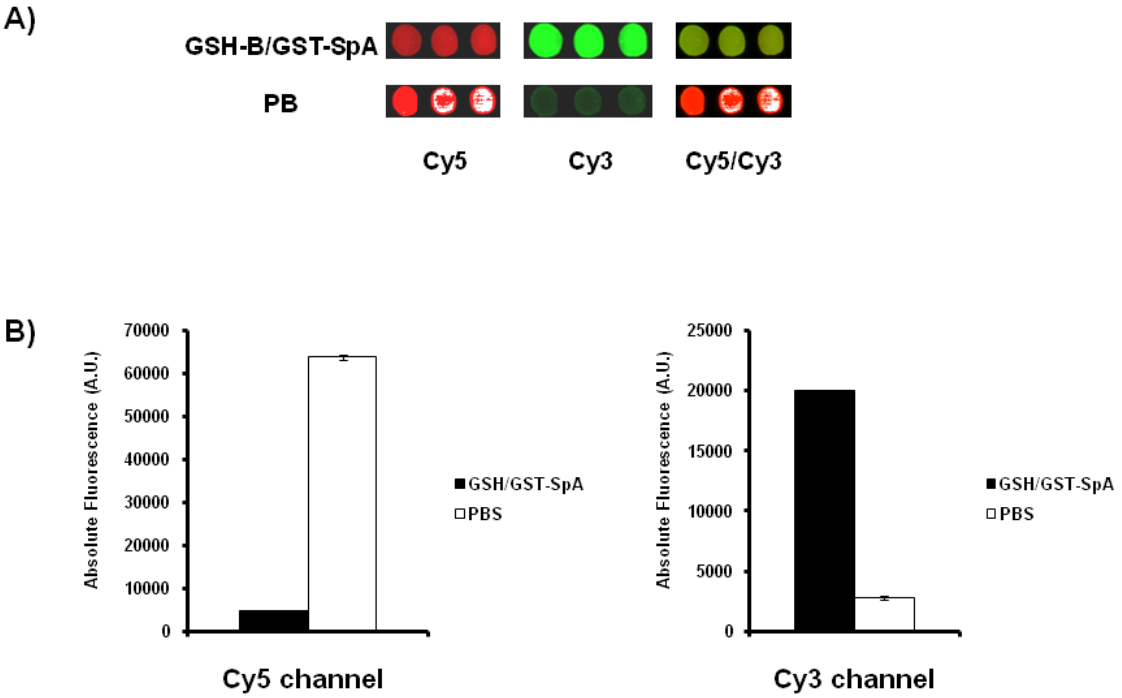

Figure 4.6 Deposition and activity of *in situ* oriented Cy5-labeled Goat α-mouse. A) Visualization of the direct deposition (Cy5 channel) and the activity of deposited antibody (Cy3 channel). A ratiometric picture is shown to aid in visualization. B) Graphical representation of the data shown in A). Cy5 channel graphs shows direct deposition of the antibody, and the Cy3 channel shows the activity of the deposited protein against Cy3-conjugated mouse IgG (50 µg). Data is representative of triplicate arrays.

As a corollary, we tested another Cy3-labeled mouse IgG to see how the oriented antibody bound to other target antibodies. After printing our *in situ* oriented and non-oriented antibodies, we probed with 50 µg of mouse α-β-tubulin (Figure 4.7). Through observing the Cy3 channel, we found that the oriented antibody showed ~10-fold more binding activity than when the antibody was randomly deposited onto the slide surface. When comparing the activity of deposited goat α-mouse IgG-Cy5 against α-actin-Cy3, we observe a 91-fold increase in activity ((relative decrease in deposition) x (relative increase in activity)). When comparing the activity of deposited goat α-mouse IgG-Cy5 against α-β-tubulin-Cy3, we observed a 130-fold increase of oriented antibody activity over non-oriented antibody.

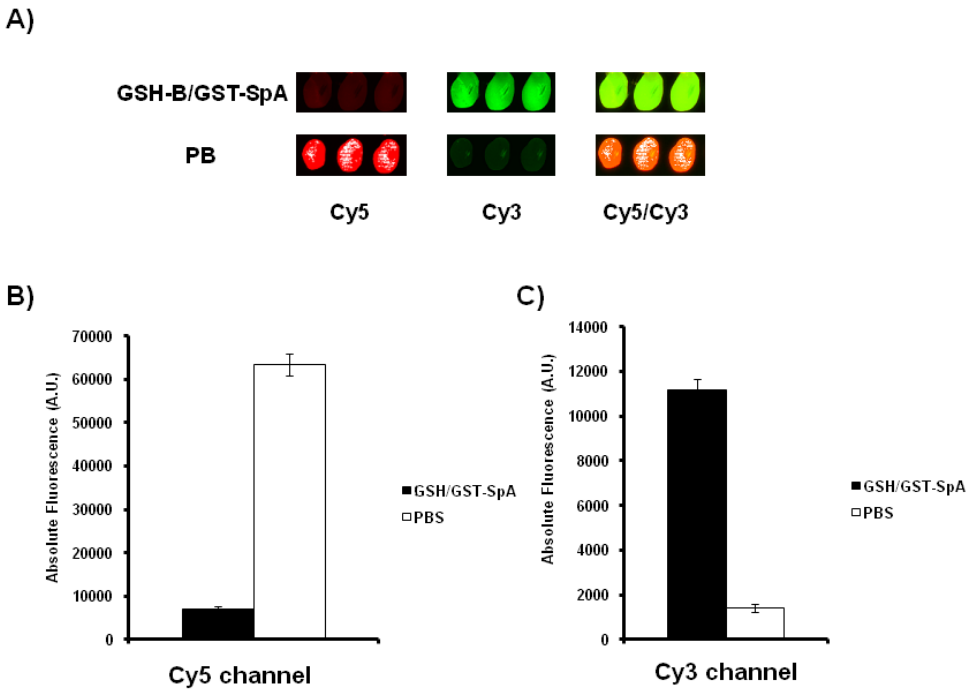

Figure 4.7 Deposition and activity of *in situ* oriented Cy5-labeled Goat α-mouse. A) Visualization of the direct deposition (Cy5 channel) and the activity of deposited antibody (Cy3 channel). A ratiometric picture is shown to aid in visualization. B) Graphical representation of the data shown in A). Cy5 channel graphs shows direct deposition of the antibody, and the Cy3 channel shows the activity of the deposited protein against Cy3-conjugated mouse IgG (50 µg, α-β-tubulin). Data is representative of triplicate arrays.

With this dual-color antibody orientation experiment, we demonstrated that we could orient the antibodies *in situ*. Our next objective was to apply this technique to the orientation of carbohydrate-binding antibodies to implement in our lectin microarrays. We chose two antibodies 7LE and 2-25LE which are Lewis A and Lewis B antigen binders, respectively. Moreover, their respective neoglycoproteins are also commercially available and can be easily converted into Cy-labeled glycoconjugates. In accordance with our optimized SpA/G (125 μg/mL) to antibody (500 μg/mL) ratios, we printed these mixtures and compared the affect of orientation by printing the antibody alone in PB. With the two antibodies tested, SpG performed the best, giving a higher median of total signals. This is not all too surprising given that these antibody-binding proteins are known to bind differently to the various classes of IgG antibodies (*21*). To combat the differences in the activity of antibody-binding proteins for different IgG subclasses, a hybrid Protein A/G could be cloned since it has been shown that it binds to all classes of IgG antibodies (*22*). After labeling the BSA conjugates with Cy3-NHS, the antibody arrays were probed with varying concentrations of Cy3-labeled Lewis A-BSA (Lewis A-Cy3, Figure 4.8) and Cy3-labeled Lewis B-BSA (Lewis B-Cy3, Figure 4.9). Testing the Lewis A-specific antibody 7LE, we observed a dramatic increase in the detection limits of the deposited antibody. Upon *in situ* orientation (GSH-B/GST-SpG), 7LE bound to Lewis A-Cy3 down to the 64 pM concentration, whereas the non-oriented 7LE (PB) could detect the same glycoprotein down to the 40 nM concentration, a remarkable increase of 625-fold in the detection limits in favor of the oriented antibody (Figure 4.8). However, the same increase in the detection limits was not observed with the Lewis B-specific antibody 2-25LE, in fact, the limit was not lowered at all (Figure 4.9). Assuming that the Cy5-labeled antibody is deposited in the same manner as these two carbohydrate-binding antibodies, we are seeing a greater increase of activity upon *in*

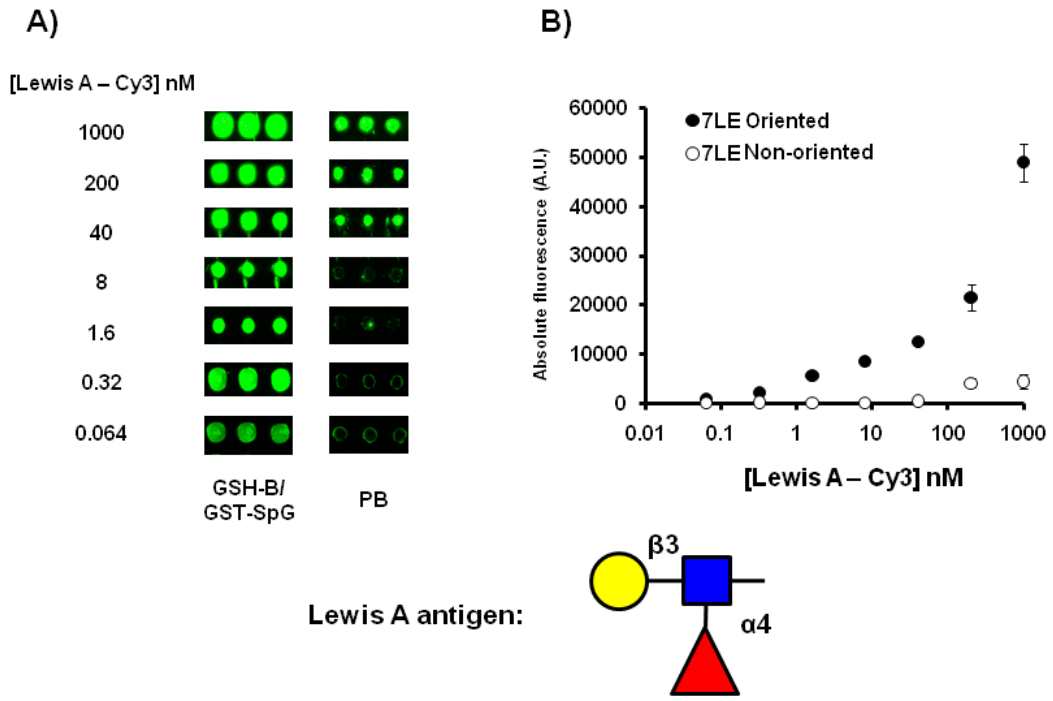

Figure 4.8 *In situ* orientation of Lewis A antibody 7LE. A) Array data showing the advantages of *in situ* oriented 7LE (GSH-B/GST-SpG) against non-oriented antibody (PB). Oriented antibody can detect down to the 64 pM concentration of glycoprotein compared to 40 nM with non-oriented 7LE, an increase of 625-fold in detection limits. B) Graphical representation of data shown in A). Data is representative of triplicate arrays and error bars represent the standard deviations from the mean.

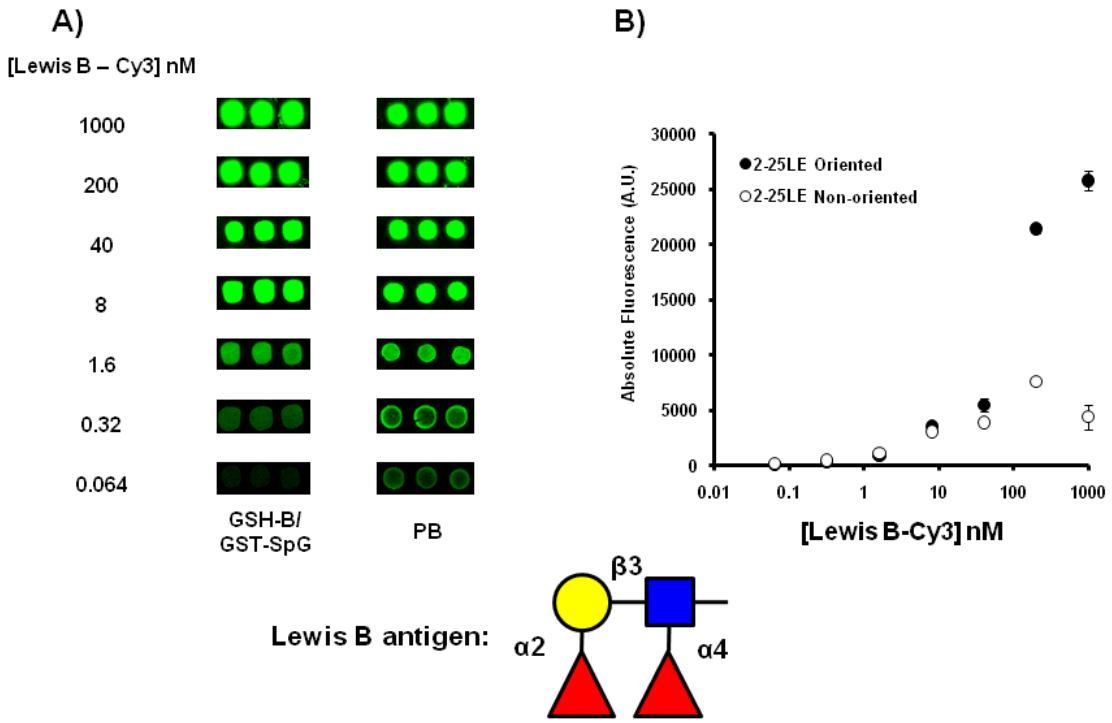

Figure 4.9 *In situ* orientation of Lewis B antibody 2-25LE. A) Array data showing the advantages of *in situ* oriented 2-25LE (GSH-B/GST-SpG) against non-oriented antibody (PB). The detection limits of the antibody is not lowered, yet the non-oriented antibody shows saturation of binding at higher concentrations whereas *in situ* oriented 2-25LE does not. B) Graphical representation of data shown in A). Data is representative of triplicate arrays and error bars represent the standard deviations from the mean.

situ orientation. Nonetheless, when observing the higher concentrations of Lewis B-Cy3, the amounts of bound ligand are higher with the *in situ* oriented antibody. This increase may be a result of saturation of the binding sites of the non-oriented 2-25LE (PB). In contrast, the binding sites of *in situ* oriented 2-25LE (GSH-B/GST-SpG) are not yet saturated, arguing for a general increase in activity against higher amounts of Lewis B-Cy3. Although we observed an increase in activity of these two antibodies, we wanted to determine if their specificity profiles changed as a result of orientation. I printed the two antibodies, 7LE and 2-25LE, both in an oriented and non-oriented manner along with the recombinant lectin PA-IIL, a fucose- and high-mannose-binding lectin (*23*). This lectin performed well as a positive control against the fucosylated Lewis antigens, but when the array was probed with RNase B-Cy3, the high-mannose containing

glycoprotein, the antibodies displayed very low activity, even when oriented (Figure 4.10). On the other hand, PA-IIL bound well to RNase B-Cy3, arguing that the oriented antibodies show increased activity against their respective glycans, while maintaining their specificity profiles.

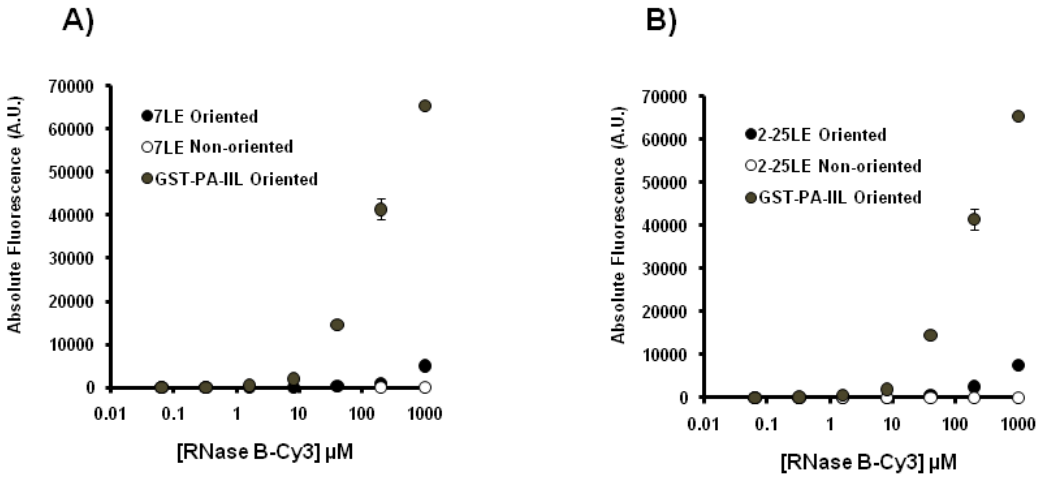

Figure 4.10 Orientation does not affect antibody activity. Activity of oriented antibodies, 7LE and 2-25LE, and PA-IIL against RNase B-Cy3. A) Activities of oriented and non-oriented 7LE and oriented PA-IIL against varying concentrations of RNase B-Cy3. B) Activities of oriented and non-oriented 2-25LE and oriented PA-IIL against varying concentrations of RNase B-Cy3. Data is representative of triplicate arrays and error bars represent the standard deviations from the mean.

4.2.2 *In situ* orientation of IgM antibodies

To orient IgM antibodies, I cloned the antibody-binding fragments of Protein L from a synthesized gene (PpL, See Appendix B for gene), and ligated the fragments into the pET41 Ek/LIC kit (Novagen). After transformation and expression, the antibody-binding proteins were purified *via* the GST-domains (Figure 4.11). To highlight the need for a novel antibody-binding protein for the orientation of IgM antibodies, we oriented all three proteins, SpA, SpG, and PpL, *in situ* and probed for IgM binding against Alexa Fluor 488 conjugated IgM isotype control (IgM-488, Figure 4.12A). When the antibody-binding proteins were probed with 2 µg of IgM-

488, we observed dramatic differences in antibody-binding activity (Figure 4.12B). This difference in antibody binding is not a great surprise if one considers the structure of this class of antibodies. When secreted from hybridoma cell lines, the IgM antibody is a multivalent antibody consisting of 5 individual IgG antibodies tethered together *via* the individual Fc domains to a central peptide called the J chain (*10*). When in the pentavalent structure, the Fc domains are buried underneath the exposed antigen-binding domains; therefore proteins which bind to Fc domains may not efficiently bind to IgM antibodies (Figure 4.12B). However, one report does mention the development of an assay that uses protein G in the detection of IgM antibodies (*24*). This was the only report I could find suggesting that SpG could be used in the orientation of IgM antibodies. Based on all of the reported binding activities, I would only expect PpL to bind IgM light chain fragments more easily, thereby orienting the IgM antibody and the remaining antigen-binding domains. Taking advantage of this structural feature will allow us to include *in situ* oriented IgM antibodies in our microarray prints.

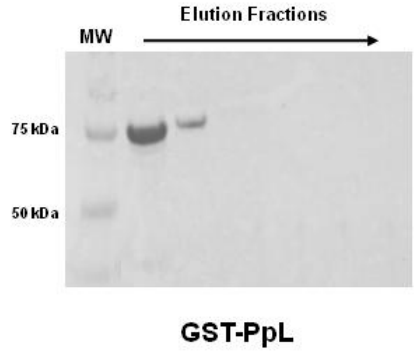

Figure 4.11 Purified GST-tagged antibody-binding protein PpL. Protein was expressed in BL21 DE3 cells and purified *via* GSH-sepharose columns. A) 10% SDS-PAGE gel analysis of PpL. GST-PpL has a molecular weight ~75 kDa.

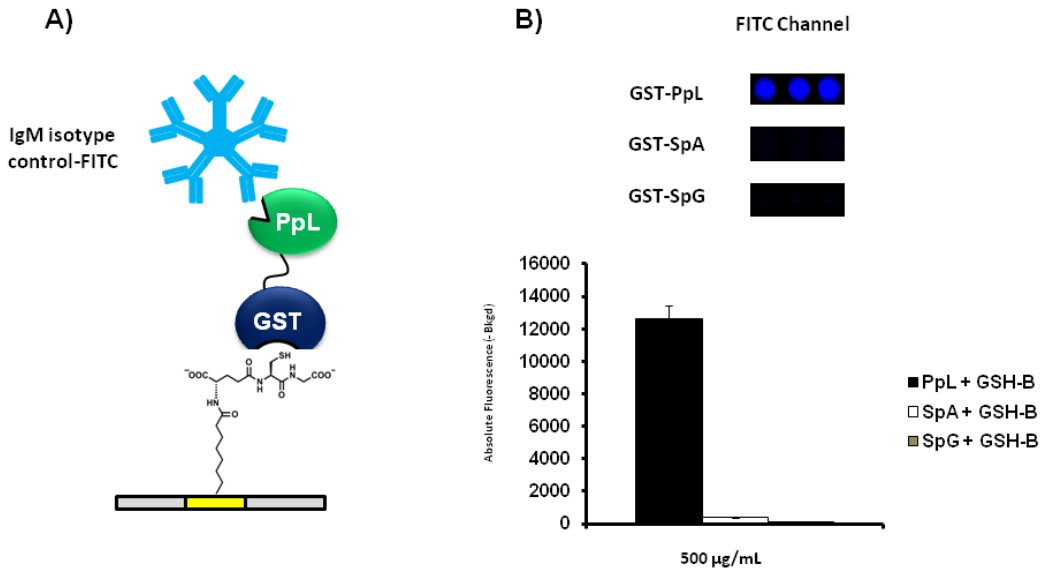

Figure 4.12 Antibody capture assay of immobilized antibody-binding proteins to IgM-488. A) Schematic of oriented GST-PpL capturing IgM-488 from solution. B) Data and graphical representation of the antibody-binding proteins activity for IgM-488.

Using our standard *in situ* orientation print buffer, we printed the Sialyl Lewis A-binding IgM antibody CA19-9 with PpL in GSH-B. As a comparison, we also printed CA19-9 in PB. The activity of the deposited antibodies was tested against varying concentrations of Cy3-labeled Sialyl Lewis A-BSA neoglycoprotein (Sialyl Lewis A-Cy3). The difference in activity of CA19-9 was mildly improved upon antibody orientation (Figure 4.13). The *in situ* oriented antibody showed a 5-fold increase in the detection limits of the deposited antibody. Although not as a significant increase like our other antibody 7LE, the oriented antibody showed consistent increases in activity, whereas the non-oriented antibody showed no significant increases in activity over every concentration of Sialyl Lewis A-Cy3 tested (Figure 4.13B). As a technical note, we observed improved spot morphology when the antibody was printed *in situ*, a common characteristic of proteins printed in this buffer (Figure 4.13A).

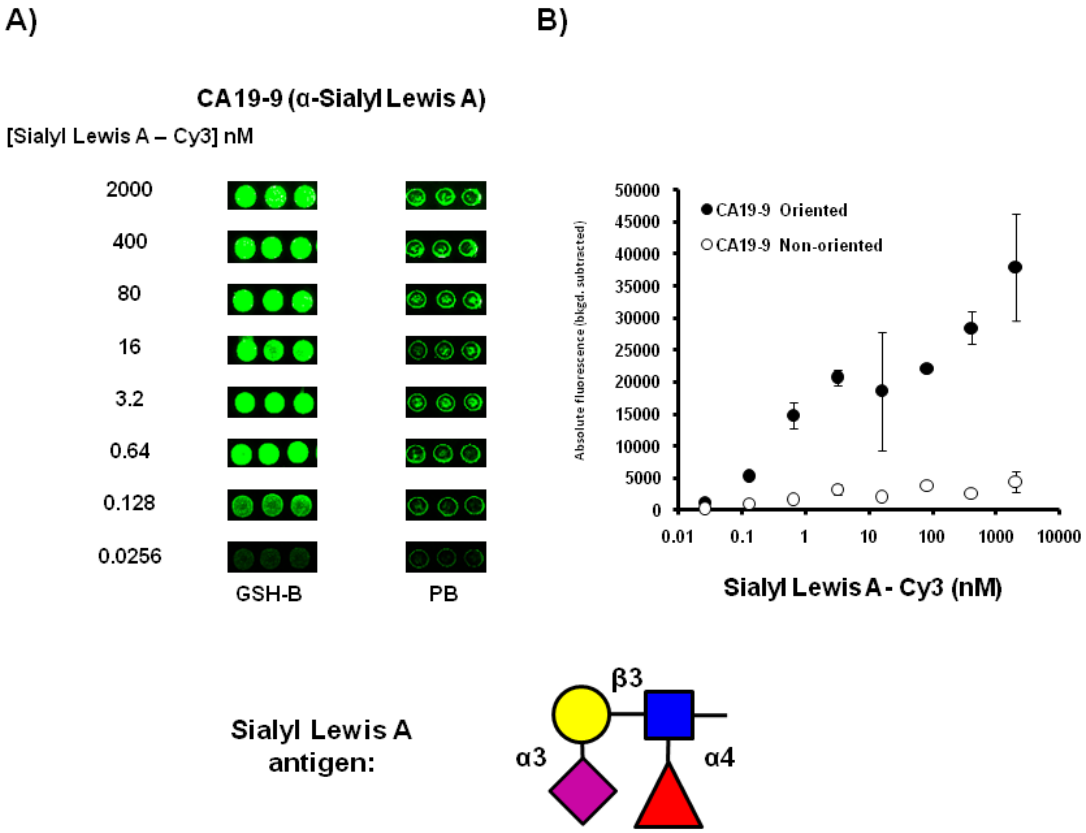

Figure 4.13 *In situ* orientation of CA19-9 (250 µg/mL). A) *In situ* oriented CA19-9 (GSH-B) was compared to non-oriented CA19-9 by probing with varying concentrations of Sialyl Lewis A-Cy3. Improved spot morphology was observed for all samples printed in GSH-B. B) Graphical representation of data shown in A). Data is representative of triplicate arrays and error bars represent the standard deviations from the mean.

Although CA19-9 worked moderately well for *in situ* orientation, three other antibodies that I tested did not perform as expected. Z5H-2 and CLCP-19B, both blood group B-binding antibodies, and 4C9, a Lewis X-binding antibody, displayed activity in an ELISA format, yet when printed with PpL in GSH-B, no activity was observed. With the high amount of GSH in our print solution, it may be that GSH reduction of disulfide bonds is fragmenting the IgM antibody resulting in lessened activity. It has been well documented that low concentrations (<1 mM) of dithiothreitol (DTT) is enough to fragment the antibody (*25,26*). In preparation and during the print process, the print solutions are incubated for at least 1 hour prior to deposition. To test the effects of GSH and DTT on antibody activity, I performed ELISA assays with

separate pre-treated mixtures of Z5H-2 and 4C9 with varying concentrations of GSH or DTT (Figure 4.14). For Z5H-2, the IC_{50} for GSH is ~ 20 mM, whereas for DTT the IC_{50} is ~1 mM (Figure 4.14A). For 4C9, the IC_{50} for GSH is ~30 mM, whereas DTT remains stronger with an IC_{50} of ~2.3 mM (Figure 4.14B). DTT was the more reactive molecule as expected, but most notably there was no visible activity at 100 mM GSH, the concentration present in GSH-B. The fact that CA19-9 antibody shows modest improvement in activity might actually be improved if the level of GSH was reduced.

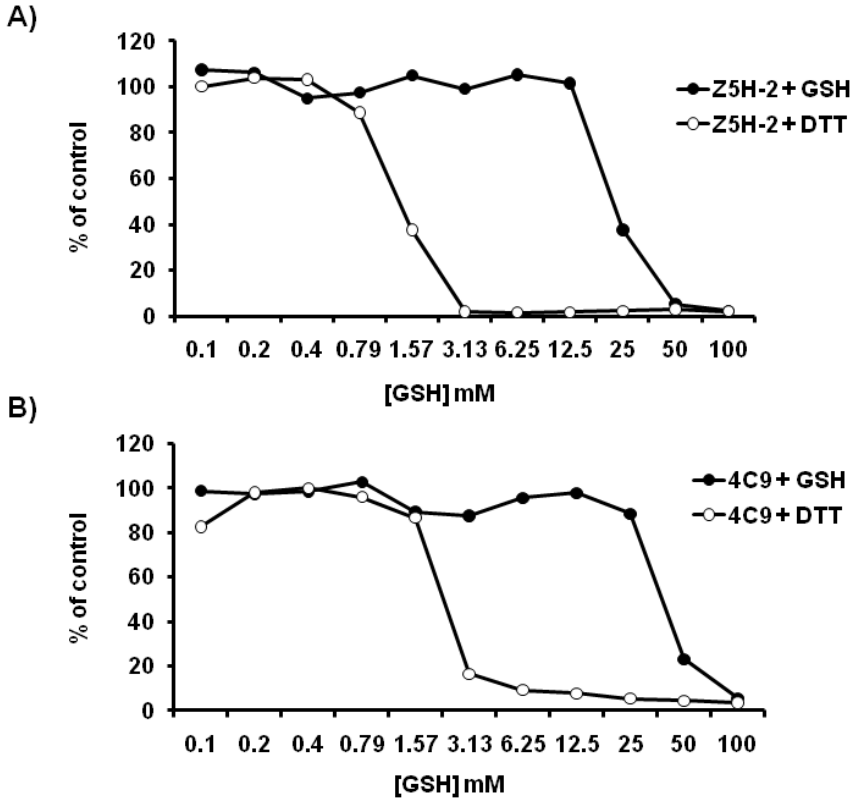

Figure 4.14 Inhibition of IgM activity by disulfide reduction using GSH and DTT. A) Effects of GSH and DTT on the activity of Z5H-2 against blood group B neoglycoprotein. DTT (IC_{50} ~1 mM) affects the antibody binding activity better than GSH (IC_{50} ~20 mM). B) Effects of GSH and DTT on the activity of 4C9 against Lewis X neoglycoprotein. DTT (IC_{50} ~2.3 mM) affects the antibody binding activity better than GSH (IC_{50} ~30 mM). Both antibodies are inhibited at the highest concentration of GSH tested, which is present in the *in situ* orientation buffer.

4.3 Conclusions

With the success of our *in situ* orientation of our GST-tagged lectins, we wanted to expand this technology to include oriented antibodies by creating a new technique in microarray technology. By printing a mixture of an IgG antibody and GST-tagged antibody-binding protein, such as SpA or SpG, we demonstrate that it is possible to orient antibodies *in situ*, thus preserving overall slide chemistry for the inclusion of non-GST-tagged proteins. We wanted to also include IgM antibodies considering this isotype constitutes ~50% of the commercially-available carbohydrate-binding antibodies. SpA and SpG are not efficient at binding to secreted IgM antibodies, so we created a GST-tagged version of PpL which was able to orient CA19-9 *in situ*. Unfortunately, only 1 of 4 antibodies tested for the *in situ* orientation of IgM antibodies showed any activity upon deposition onto the slide surface. After further experiments, we believe that this loss of activity can be attributed to the high concentration of GSH in the print buffer. The disulfide bridges holding the IgM antibody seem to be very reactive toward this high concentration of reducing agent. This does not seem to be the case in IgG antibodies, but further work will be needed to elucidate the effects of GSH on this class of antibodies. As an alternative, current work in the Mahal lab is looking into expanding the *in situ* orientation idea to multi-His-tagged (*27*) proteins using an excess of Ni^{2+}-NTA ligand in the print buffer. It may be possible to use multi-His-tagged antibody-binding proteins to orient the reducing-agent susceptible antibodies.

4.4 Materials and Methods

4.4.1 Cloning, expression, and purification of GST-tagged SpA, SpG, and PpL

Protein A (SpA) was cloned from *Staphylacoccus aureus subsp. aureus*, strain Mu3 (ATCC #700698D-5). Protein G (SpG) was cloned from a synthetic gene that was codon optimized for *Escherichia coli* (see Appendix B). Protein L was cloned from a synthetic gene that was codon optimized for *E. coli* expression (see Appendix B). Protein A primers; Forward: 5' GAC GAC GAC AAG ATG GCT GAT AAC AAT TTC AAC AAA GAA CAA CAA AAT GC 3', Reverse: 5' GAG GAG AAG CCC GGT TTA AGC ATC GTT TAG CTT TTT AGC 3'. Protein G primers; Forward: 5' GAC GAC GAC AAG ATG ACG TAT AAA CTG ATC CTG 3', Reverse: 5' GAG GAG AAG CCC GGT CTA TTC CGT CAC GGT GAA TG 3'. Protein L primers; Forward: 5' GAC GAC GAC AAG ATG AAA GAG GAA ACC CCG 3', Reverse: 5'GAG GAG AAG CCC GGT TTA ACC GGC GAA ACG AAT G 3'. PCR reactions were performed with the Hot Start DNA Polymerase (New England Biolabs, #F120S) and reactions were tailored to product specifications. PCR products were treated and ligated into the pET-41 Ek/LIC kit (Novagen, #71017-3) according to the manufacturer's directions. DNA was then transformed into electrocompetent NovaBlue Gigasingles (Novagen, #71227), grown, and purified DNA was obtained using the Qiaprep Mini Kit (Qiagen, #27106), and sequenced. Positive sequences were transformed into electrocompetent BL21(DE3) cells, and grown on LB-Agar supplemented with kanamycin (30 μg/mL). Colonies were picked for overnight growth in LB, and were then inoculated into 900 mL of LB supplemented with kanamycin (30 μg/mL). Cultures were grown to an OD_{600} 0.7 – 1.0, then induced with 0.2 mM IPTG and grown for 3 hr at 37°C, shaking at 250 rpm. Cells were pelleted (3000 x g, 15 min) and resuspended in 40 mL of lysis buffer (PBS + 0.2% Triton-X10) with protease inhibitor mix. Lysozyme (~1 mg/mL) was

added and mixed on ice for 30 min. DNase I (New England Biolabs #M0303, ~5 units/mL of lysate) was added and mixed for 10 min on ice. Mixture was then centrifuged at 30,000 x g for 30 min. The supernatant was then loaded onto a 1 mL GSH-sepharose column (GE Healthcare, # 17-0756-01), washed with 10 mL of PBS, and eluted with 10 mM GSH in 50 mM Tris, pH 8.0, and collecting 1 mL fractions. Purification was analyzed by SDS-PAGE (Figure 4.3 and 4.11), and fractions containing sample were pooled and dialyzed against PBS. Aliquots were prepared and snap frozen in liquid N_2 and stored at -80°C until needed.

4.4.2 *In situ* oriented antibody microarray

For the oriented antibody array using the IgG antibodies, the antibodies and GST-SpA or -SpG were printed in ~1.1 : 1 molar ratio (0.5 mg/ mL Ab: 0.125 mg/mL GST-SpA/SpG) in GSH-B. For standard printing, antibodies were printed in PBS. For dual color antibody printing, Cy5-labeled Goat α-Mouse IgG antibody (0.5 mg/mL, Abcam #ab6563) and GST-SpA (0.125 mg/mL) were diluted in GSH-B or PB. Samples were printed as described above, and the array was hybridized for 2 hr with either Cy3-labeled Mouse IgG α-Actin (2 µg/mL in 0.005% PBS-T, Abcam #ab11004) or Cy3-labeled Mouse IgG α-β-Tubulin (2 µg/mL in 0.005% PBS-T, Abcam #ab11309). For the analysis of the carbohydrate-binding antibodies, Lewis A-specific antibody (0.5 mg/mL, 7LE, Abcam #ab3967) or Lewis B-specific antibody (0.5 mg/mL, 2-25LE, Abcam #ab3968) and GST-SpG (0.125 mg/mL) were dissolved in GSH-B or PB. Samples were printed as described above, and were hybridized in 0.005% PBS-T with either Cy3-labeled LNFP II-BSA (Lewis A neoglycoprotein, Dextra, UK #NGP0501) or Cy3-labeled LNDFHI-BSA (Lewis B neoglycoprotein, Dextra #NGP0601). The neoglycoproteins were labeled with Cy3-NHS (GE Healthcare, #PA13104) prior to hybridization.

For the orientation of CA19-9, the antibody (0.25 mg/mL, Abcam #ab3982) and GST-PpL (0.125 mg/mL) were diluted in GSH-B, or CA19-9 was printed alone in PB. Samples were printed as described above, and the array was hybridized for 2 hr with varying concentrations of Cy3-labeled Sialyl Lewis A-BSA conjugate (2000, 400, 80, 16, 3.2, 0.64, 0.128, and 0.0256 nM) diluted into 0.005% PBS-T. The neoglycoproteins were labeled with Cy3-NHS (GE Healthcare, #PA13104) prior to hybridization.

4.4.3 ELISA activity assays

Enzyme-linked immunosorbent assays (ELISAs) were performed as previously described (*28*) with specific adaptations for the detection of antibodies. Briefly, 100 µL of blood group B-BSA (Dextra UK), for Z5H-2 and CLCP-19B, and Lewis X-BSA (Dextra UK), for 4C9, was dissolved in PBS (5 ng/µL) and added to 96-well microtiter plates (Grenier Bio-one, #655061, Monroe, NC). Following incubation (overnight at 4°C), the plates were washed five times with 0.05% PBST (PBS and 0.05% Tween 20). Wells were then blocked with PBS containing 5% BSA for 1 h at RT and again washed five times with 0.05% PBST. Upon drying, 50 µL of Z5H-2 (1:5000), CLCP-19B (1:2500), and 4C9 (1:5000) in 1% BSA in 0.05% PBST were added to separate wells and incubated for 1 hr. For the inhibition experiments, antibody mixtures were pre-incubated for 1 hr with varying amounts of GSH or DTT, and then the samples added to the wells and incubated for 1 hr. In all experiments, wells containing buffer alone were used to measure the background (noise) of the plate. After the incubation times, wells were washed five times with 0.05% PBST and lectin binding was detected with anti-mu chain-horseradish peroxidase (HRP) conjugated antibody (50 µL, 1 : 1000 in 0.05% PBST, 1 h, RT, Novus Biologicals No. NB600-393, Littleton, CO). Wells were then washed 5 times with PBST and HRP activity was detected using *o*-phenylenediamine dihydrochloride (OPD Pierce No. 34 005,

Rockford, IL). In brief, OPD solution (100 mL, 0.4 mg/mL in 0.1 M phosphate/citrate, pH 5.0 containing 0.004% H_2O_2) was added to each well. After 30 min, the enzymatic reaction was stopped by the addition of 50 mL of 2.5 M H_2SO_4. The absorbance at 492 nm was read using a Synergy HT microplate reader (BIO-TEK, Winooski, VT). A reference wavelength of 620 nm was subtracted from these values to account for non-specific absorbance. For data analysis, the antibody binding activity was defined as the signal to noise (S/N): absorbance of samples divided by the average background absorbance values. To determine inhibition of reducing agents, percent of control (% control) was calculated: (the absorbance of the sample in the presence of inhibitor/absorbance of uninhibited sample)/100. Microsoft Excel 2007 software was used for statistical analysis, curve fitting, and to generate graphs and tables.

4.5 References

1. Wingren, C., and Borrebaeck, C. A. K. (2004) High-throughput proteomics using antibody microarrays. *Expert. Rev. Proteomics 1*, 355 – 364.
2. Wingren, C., and Borrebaeck, C. A. K. (2006) Antibody microarrays: current status and key technological advances. *Omics 10*, 411 – 427.
3. Borrebaeck, C. A. K., and Wingren, C. (2009) Design of high-density antibody microarrays of disease proteomics: Key technological issues. *J. Proteomics 72*, 928 – 935.
4. Haab, B. B., Dunham, M. J., and Brown, P. O. (2001) Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Gen. Biol. 2*, 1 – 13.
5. Haab, B. B. (2003) Methods and applications of antibody microarrays in cancer research. *Proteomics 3*, 2116 – 2122.
6. Hucknall, A., Kim, D.-H., Rangarajan, S., Hill, R. T., Reichert, W. M., and Chilkoti, A. (2009) Simple fabrication of antibody microarrays on nonfouling polymer brushes with femtomolar sensitivity for protein analytes in serum and blood. *Adv. Mat. 21*, 1968 – 1971.
7. Taeusch, H. W., Ballard, R. A., Gleason, C. A., and Avery, M. E. (2005) Avery's diseases of the newborn. *Elsevier Health Sciences, 8th Ed. Ch. 35*, 454.
8. Butch, A.W., Macke, K. A., Scott, M. G., Inkster, M., and Nahm, M. H. (1989) Mitogen-induced human IgG subclass expression. II. IgG1 and IgG3 subclasses are preferentially stimulated by a combination of *Staphylococcus aureus* Cowan I and pokeweed mitogen. *Hum. Immunol. 24*, 207 – 218.

9. Cathou, R. E., and O'Konski, C. T. (1970) A transient electric birefringence study of the structure of specific IgG antibody. *J. Mol. Biol. 48*, 125 – 131.

10. Plaut, A. G., and Tomasi, T. B. Jr. (1970) Immunoglobulin M: pentameric Fcmu fragments released by trypsin at higher temperatures. *Proc. Natl. Acad. Sci. 65*, 318 – 322.

11. Wiersma, E. J., Collins, C., Fazel, S., and Shulman, M. J. (1998) Structural and functional analysis of J chain-deficient IgM. *J. Immunol. 160*, 5979 – 5989.

12. El-Sayed, A., Alber, J., Lammler, C., Abdulmawjood, A., Zschock, M., and Castaneda, V. H. (2006) Comparative sequence analysis of *spa* gene of *Staphylococcus aureus* isolated from bovine mastitis: characterization of an unusual *spa* variant. *J. Dairy Res. 73*, 322 – 327.

13. Guss, B., Eliasson, M., Olsson, A., Uhlen, M., Frej, A.-K., Jornvall, H., Flock, J.-I., and Lindberg, M. (1986) Structure of the IgG-binding regions of streptococcal protein G. *EMBO J. 5*, 1567 – 1575.

14. Derrick, J. P., and Wigley, D. B. (1992) Crystal structure of a streptococcal protein G domain bound to an Fab fragment. *Nature 359*, 752 – 754.

15. Wang, H., Liu, Y., Yang, Y., Deng, T., Shen, G., and Yu, R. (2004) A protein A-based orientation-controlled immobilization strategy for antibodies using nanometer-sized gold particles and plasma-polymerized film. *Anal. Biochem. 324*, 219 – 226.

16. Matson, R. S., Milton, R. C., Rampal, J. B., Chan, T. S., and Cress, M. C. (2005) Overprint immunoassay using protein A microarrays. *Meth. Mol. Biol. 382*, 273 – 286.

17. Vareiro, M. M. L. M., Liu, J., Knoll, W., Zak, K., Williams, D., and Jenkins, A. T. A. (2005) Surface plasmon fluorescence measurements of human chorionic gonadotrophin:

Role of antibody orientation in obtaining enhanced sensitivity and limit of detection. *Anal. Chem. 77*, 2426 – 2431.

18. Chen, M. L., Adak, A. K., Yeh, N. C., Yang, W. B., Chuang, Y. J., Wong, C. H., Hwang, K. C., Hwu, J. R., Hsieh, S. L., and Lin, C. C. (2008) Fabrication of an oriented Fc-fused lectin microarray through boronate formation. *Angew. Chem. Int. Ed. Engl. 47*, 8627 – 8630.

19. Ha, T. H., Jung, S. O., Lee, J. M., Lee, K. Y., Lee, Y., Park, J. S., and Chung, B. H. (2007) Oriented immobilization of antibodies with GST-fused multiple Fc-specific B-domains on a gold surface. *Anal. Chem. 79*, 546 – 556.

20. Kastern, W., Sjobring, U., and Bjorck, L. (1992) Structure of *Peptostreptococcal* protein L and identification of a repeated immunoglobulin light chain-binding domain. *J. Biol. Chem. 267*, 12820 – 12825.

21. Harlow, E., and Lane, D. eds (1988) Antibodies: A laboratory manual. *Cold Spring Harbor Laboratory, N. Y.*, 617 – 618.

22. Eliasson, M., Olsson, A., Palmcrantz, E., Wiberg, K., Inganas, M., Guss, B., Lindberg, M., and Uhlen, M. (1988) Chimeric IgG-binding receptors engineered from staphylococcal protein A and streptococcal protein G. *J. Biol. Chem. 263*, 4323 – 4327.

23. Wu, A. M., Gong, Y.-P., Li, C.-C., and Gilboa-Garber, N. (2010) Duality of the carbohydrate-recognition system of *Pseudomonas aeruginosa*-II lectin (PA-IIL). *FEBS Lett. 584*, 2371 – 2375.

24. Zatta, P. F. (1996) A new bioluminescent assay for studies of protein G and protein A binding to IgG and IgM. *J. Biochem. Biophys. Meth. 32*, 7 – 13.

25. Beale, D., and Feinstein A. (1969) Studies on the reduction of a Human 19s immunoglobulin M. *Biochem. J. 112*, 187 – 194.

26. Kownatzki, E. (1973) Disulfide bonds of human IgM: Differential sensitivity to reductive cleavage. *Scand. J. Immunol. 2*, 433 – 437.

27. Fischer, M., Leech, A. P., and Hubbard, R. E. (2011) Comparative assessment of different histidine-tags for immobilization of protein onto surface plasmon resonance sesnsorchips. *Anal. Chem. 83*, 1800 – 1807.

28. Hsu, K.-L., Gildersleeve, J. C., and Mahal, L. K. (2008) A simple strategy for the creation of a recombinant lectin microarray. *Mol. BioSys. 4*, 654 – 662.

Chapter 5: Designed lectin variants as tools in microarray technology

5.1 Introduction

Within our set of recombinant lectins, we have two classes of bacterial lectins (Figure 5.1): fimbrial and tetrameric, soluble adhesins. The fimbrial adhesins include GafD (*1*), PapGII (*2*) and PapGIII (*3*), and FimH (*4*), although in our lab we have yet to obtain FimH lectin that binds to mannosylated conjugates, a continuing issue to be resolved in the group. Fimbrial pilin structures are produced in Gram-negative bacteria, and in this case arise from certain strains of *Escherichia coli* (*5*). In the examples cited above, the carbohydrate- binding proteins are located at the terminus of the pilin domain, and in *E. coli*, there are typically over 200 fimbrial domains on the cell surface (*6*). Although structurally similar, the fimbrial lectins bind a diverse set of glycans, arguing for divergent evolution from a common scaffold (*7*). GafD is a terminal β-GlcNAc-binding lectin derived from *E. coli* strains isolated from bovine intestine (*1*). PapGII and III are two alleles that have high sequence homology, but which are found in species-dependent uropathogenic bacteria (*8*). PapGII-containing *E. coli* strains effectively infect human and mouse tissues and PapGIII cannot, andyet PapGIII-containing strains can infect canine tissue (*2-3*). These differences are based on the biosynthesis of their corresponding glycolipids. The PapG adhesion family bind to the globoside glycan family; PapGII binds to globotetratose (GbO4, GalNAcβ1-3Galα1-4Galβ1-4Glc-Cer) and PapGIII binds to globopentaose (GbO5, GalNAcα1-3GalNAcβ1-3Galα1-4Galβ1-4Glc-Cer) (*8*). The other known allele of this family, PapGI, was not developed prior to this work, but I will show that it too can be cloned, expressed and retain its known binding activity for globotriose (GbO3, 3Galα1-4Galβ1-4Glc-Cer).

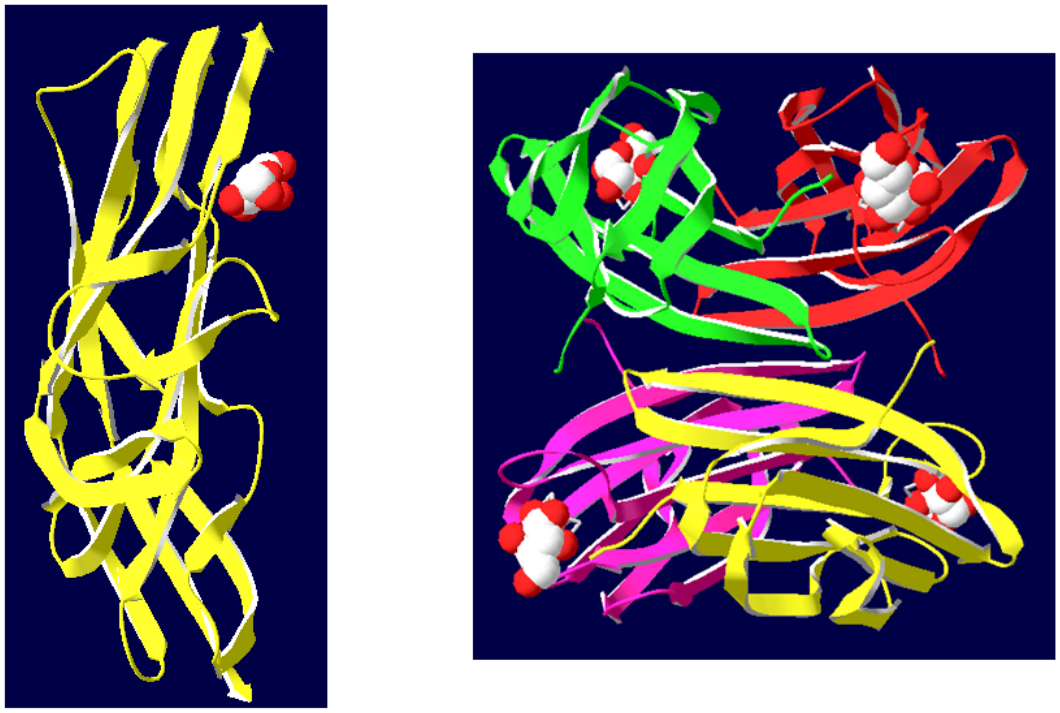

Figure 5.1 Examples of the structures of two types of recombinant lectins. GafD (PDB 1OIO), a pilin adhesin, lies at the tip of a column of pilin proteins. The structure of these pilin adhesins adopts a β-barrel jelly roll fold and has been compared to the fold of immunoglobulin-type folds. PA-IL (PDB: 2VXJ) is a tetrameric, soluble adhesion which is secreted by *Pseudomonas aureuginosa* to mediate bacterial interactions. Crystal structures were isolated with their corresponding carbohydrates.

The second type of bacterial lectins is the Ca^{2+}-dependent, tetrameric, soluble adhesins. The opportunistic human pathogen *Psuedomonas aeruginosa* secretes two bacterial lectins, PA-IL and PA-IIL (also called LecA and LecB) (*9*). PA-IL is a galactophilic lectin with specificity for terminal α-galactose residues (*10*). PA-IIL is primarily a fucose-binding lectin that has the ability to bind mannose epitopes, but to a lower degree (*11*). The production of these lectins is regulated through quorum sensing (*12*), and as such, has been linked to biofilm formation. In a related species, the plant pathogen *Ralstonia solanacerum* produces two lectins as well, RS-IL and RS-IIL, both lectins being orthologous to the PA-IIL lectin (*13*). RS-IL is a fucose-binding lectin, and RS-IIL is mainly a mannose-binding lectin, but does show some binding to

fucosylated epitopes to a lower degree (*13*). The lectins are homologous, i.e. RS-IIL and PA-IIL share a common binding motif in the carbohydrate-binding domain. Interestingly, the binding domains of these lectins differ in a single amino acid which can shift binding preferences from fucose to mannose, or vice versa (*14*). The first section of this chapter will focus on the lectins mentioned above and the creation of simple binding mutants that can act as negative controls on a given lectin microarray. The second half of this chapter will focus on my efforts to evolve of one of these lectins, GafD.

5.2 Results and Discussion

5.2.1 Controls in protein microarray technology

As discussed earlier in this dissertation, recombinant lectins have several advantages over the more commonly used native plant lectins (*15*). One advantage lies in the ease of genetic manipulation, wherein we could mutate the binding domain to bind a different antigen. In previous work developed by the lab, Hsu et al. generated a subset of GST-tagged lectins (*15*) which I used to orient on a solid support (Chapters 2 and 3). Among these lectins was a mutant of GafD (GafD-m), a single point mutation in the binding site (D88L) that showed an ~80% reduction in binding to GlcNAc (*16*). In the original work on the mutant, how the activity was assessed was not described. Nonetheless when subcloned into the pET41 vector, GafD-m displayed no significant binding to glycoproteins, although inhibitable activity was observed when tested against membrane-bound glycans (data not shown) (*15*). This binding mutant was useful in determining real versus non-specific interactions; therefore we wondered whether we could generate a binding mutant for each recombinant lectin. Non-binding mutants of the lectins would act as negative controls for recombinant lectins that show binding activity in a microarray.

Given the volume of data collected in microarray technology, controls are a crucial aspect of this technique. In gene microarrays, negative controls, which are specific for a set of experiments, provide background binding data which the positive signals can be compared against (*17*). In lectin microarray technology, such negative control lectins do not exist. The specificity of lectin-carbohydrate interactions are determined by the disruption of lectin activity with inhibiting monosaccharides. This method is imprecise because most lectins bind to di- and trisaccharides, thus inhibitions with monosaccharides are not very effective with strong interactions. Furthermore, inhibitions require the use of multiple arrays, both limiting the number of samples tested and increasing the amount of sample used. Production of GafD-m opened up the possibility of a more precise set of negative controls in the form of a matched pair of non-binding mutants and wild-type recombinant lectins.

5.2.2 Non-binding mutants of the pilin adhesins GafD, PapGI, II, and III

As discussed earlier, GafD is a fimbrial protein which binds to terminal β-linked *N*-acetylglucosamine (β-GlcNAc) epitopes (*18*). In earlier work, we demonstrated that the GafD adhesin can be expressed as a GST-fusion protein and still bind to GlcNAc-containing residues on a glycan microarray (*15*). In this work, we also showed that a binding mutant (GafD-m, GafD D88L) shows less binding to the same glycan epitopes when compared at identical concentrations (*15*). The creation of this GafD variant provides a significant control for lectin microarrays. However, the effectiveness of this mutant was reported to reduce binding activity by 80% (*16*). When probed against samples containing high amounts of terminal GlcNAc epitopes, residual activity can arise since binding is reduced and not absent. This issue prompted us to design another GafD variant based on a solved crystal structure with β-GlcNAc nestled in the binding domain (*16*). We generated several variants of GafD by mutating residues in and

around the glycan-binding domain to alanine (Figure 5.3A). The Trp109 residue is common in other pilin adhesins FimH and PapG (*7*), and aromatic side chains are known to create key hydrophobic binding contacts in other plant lectins (*19*). Cys110 is not directly involved in the binding domain, yet forms a stable, internal disulfide bridge which may influence the shape of the binding pocket. The hydroxyl residue of Thr117 makes formative hydrogen bonding contacts with the N-acetyl group at the C2 position of the glycan (*16*). Because of this interaction, we generated a double mutant (D88L + T117A) which we hoped would solve issues of any residual binding from GafD D88L.

To screen many samples at once, we developed a glycoprotein microarray, the inverse of a lectin microarray, capable of detecting the binding of our recombinant lectins to known glycan epitopes (Figure 5.2). A variety of glycoproteins and neoglycoproteins were purchased, diluted to 0.5 mg/mL in phosphate buffered saline and printed onto NHS-activated slides (Table 5.1). The glycoprotein microarray was developed by another member of the lab, Linlin Wang, and was originally developed to screen new carbohydrate-binding proteins of unknown specificity. However, in the development process, the original recombinant lectins were used to test the effectiveness of the array and since the method proved to be robust, we decided to test our binding variants in concert with the wild type recombinant lectins. In brief, the recombinant lectins were diluted into 0.005% PBS-Tween with 1% BSA and hybridized to the array. After washing away unbound lectin, the bound lectins were detected using an Alexa Fluor 488-conjugated α-GST antibody (α-GST-488). The data shown in the following sections is shown as a heat map, where we took the $\log_2$ of the background subtracted values and applied a color scheme to the values. In our standard lectin microarrays, we typically use a fluorescence value of

1000 arbitrary units (A.U.) and above to determine a positive signal, and used the value of 9.97 as a cutoff point for activity ($Log_2 1000 = 9.97$).

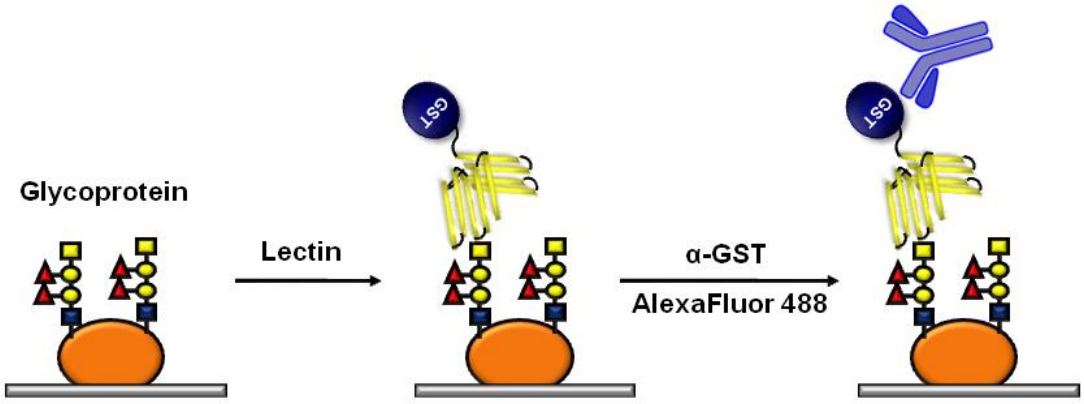

Figure 5.2 Schematic of glycoprotein microarray. The glycoproteins are immobilized onto an NHS-activated surface. The GST-tagged lectins are the incubated with the array and are washed away. Lectin-binding interactions were then probed with alexa fluor 488-conjugated α-GST antibody.

Glycan Description	Abbreviation	Supplier
α1-Acid Glycoprotein from bovine plasma 99%	AGP	Sigma
Asialofetuin from fetal calf serum Type I	ASF	Sigma
Bovine submaxillary Mucin	BSM	Worthington
Fetuin from fetal calf serum	Fet	Sigma
Ovalbumin	Ova	Worthington
Ovomucoid, Trypsin Inhbitor	Ovo	Worthington
Ribonuclease B from bovine pancreas	RNase B	Sigma
Thyroglobulin from bovine thyroid	Thy	Sigma
Transferrin human	TrF	Sigma
2'Fucosyllactose-BSA(2'FL-BSA) (3 atom spacer)	2'FL	DextraUK
3'-Sialyl Lewisx-BSA(14 atom spacer)	SLe^x	DextraUK
3'-Sialyl-3-fucosyllactose-BSA(3 atom spacer)	SFLac	DextraUK
3'-Sialyl-N-acetyllactosamine-BSA(14 atom spacer)	SLacNAc	DextraUK
6'Sialyllactose-APD-HSA	SLac	Glycotech
Acetylgalactosamine-BSA (14 atom spacer)	GalNAc	DextraUK
β1-4-Galactosyl-Galactose-BSA(3 atom spacer)	Galb1-4Gal	DextraUK
Blood group A-BSA (6 atom spacer)	BGA	DextraUK
Blood group B-BSA (6 atom spacer)	BGB	DextraUK
D-Galactose-BSA (14 atom spacer)	Gal	DextraUK
D-Mannose-BSA (14 atom spacer)	Man	DextraUK
Gala1-3Galb1-4Glc-HSA(3 atom spacer)	iGbO3	DextraUK
Gala1-3Galb1-4GlcNAc-BSA(14 atom spacer)	α-Gal	DextraUK
Gala1-3Gal-BSA(14 atom spacer)	Gala1-3Gal	DextraUK
Globotriose-HSA (3 atom spacer)	GbO3	DextraUK
Lacto-N-fucopentaose II-BSA(LNFP II-BSA) (3 atom spacer)	Le^a	DextraUK
Lewisx-BSA(3 atom spacer)	Le^x	DextraUK
L-Fucose-BSA (14 atom spacer)	Fuc	DextraUK
LNF I-APD-HSA	Le^d	Glycotech
LNF III-APD-HSA	Le^x	Glycotech
N-Acetylglucosamine-BSA (14 atom spacer)	GlcNAc	DextraUK
N-Acetyllactosamine-BSA(14 atom spacer)	LacNAc	DextraUK
Sialyl-LNF V-APD-HSA	SLNFV	Glycotech

Table 5.1: List of glycoproteins and neoglycoproteins printed in glycan microarray.

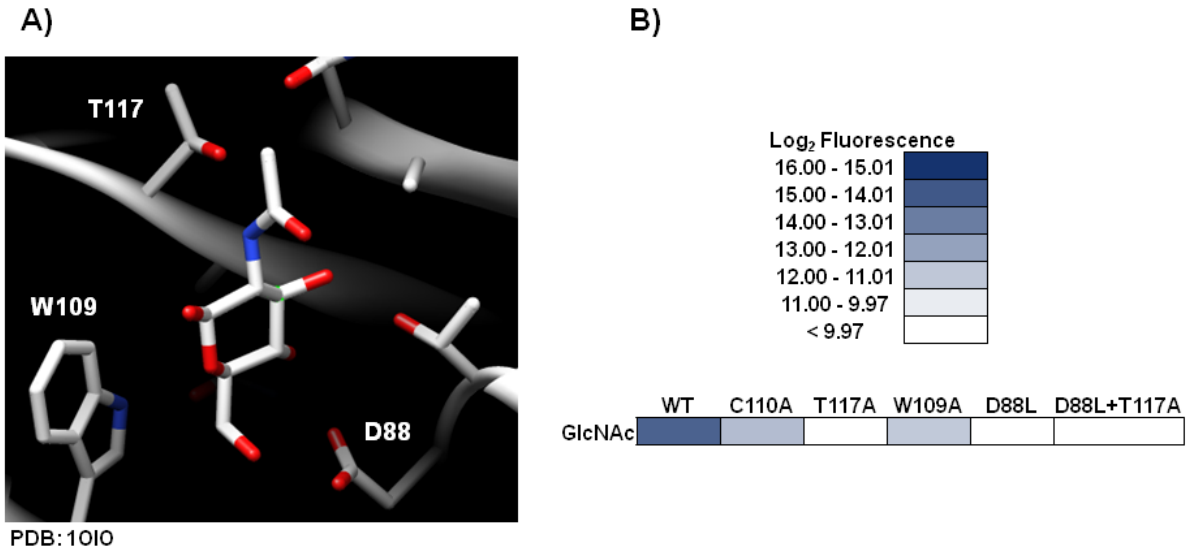

Figure 5.3 Mutational analysis of GafD lectin. A) Visualization of the crystal structure of GafD with GlcNAc in the binding site is shown the left. B) Heat map representation of GafD wt and variant binding to GlcNAc-BSA. Of all the mutants, the double mutant D88L + T117 showed the lowest overall activity.

As expected, wild-type GafD (GafD WT) bound to GlcNAc-BSA on our glycoprotein microarray (19,140 A. U.) (Figure 5.3B). Contrary to our previous work with GafD, the lectin showed no significant binding to chicken egg ovalbumin possibly due to the use of a different lot of glycoprotein, which can vary in glycosylation. Our GafD C110A variant displayed an 85% reduction in binding to GlcNAc-BSA when compared to GafD WT, arguing that the cysteine residue provides some structural integrity of the binding domain (Figure 5.3B). It is unclear whether the disulfide bridge is necessary given that when exposed to high amounts of GSH, the GafD lectins is still active. The GafD W109A variant activity was reduced to 86% of the WT binding to GlcNAc-BSA, illustrating the importance of hydrophobic pockets in glycan-binding interactions. By far, the best non-binding mutants were GafD D88L, GafD T117A, and GafD D88L + T117A. With respect to GlcNAc-BSA, activity for all three variants was reduced to ~500 A.U. (~98% reduction in binding). We are currently examining the activities of these variants against more complex carbohydrate samples.

PapG alleles class I-III are associated with infections of the urinary tract (*20*). The PapG adhesins recognize Galα1-4Gal core of the globoside lipids. In previous work, we demonstrated that PapGII and PapGIII are expressed and bind to their respective ligands in a recombinant format (*15*). To further study the globoside class of glycolipids, we added PapGI to our panel of recombinant lectins by cloning it from *Escherichia coli* strain J96 (gift from Prof. Rodney Welch, Univ. Mad.-Wisc.) (*21*). It has been shown that the three alleles have differing specificities to the globosides. PapGI binds to the trisaccharide GbO3 (Pk antigen), PapGII binds to the tetrasaccharide GbO4, and PapGIII binds to GbO5 (Forssman antigen) (*8*). A sequence alignment of the three adhesins and a crystal structure of PapGII with GbO4, revealed the critical

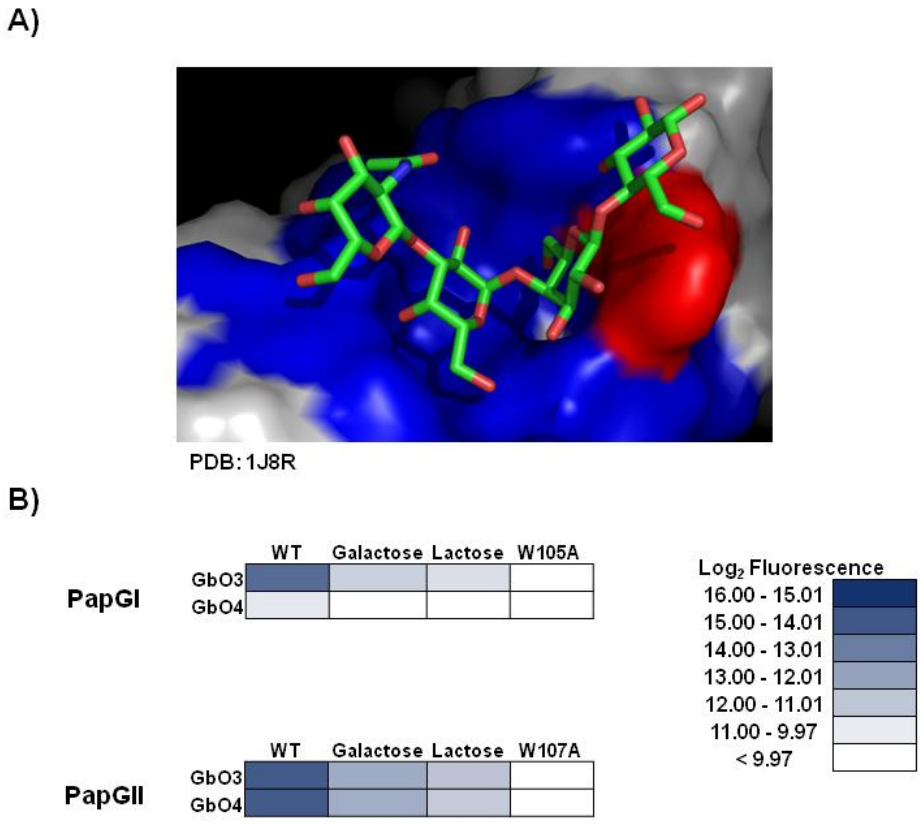

Figure 5.4 Mutational analyses of the PapG lectins. A) Visualization of the crystal structure of PapGII with GbO4. Residues with contact with the ligand are shown in blue and the W109 residue is highlighted in red. B) Heat map representation of the binding of wt PapG lectin, with inhibitions, compared to the tryptophan mutants.

residues for glycan-binding (Figure 5.4A). Based on previous work on the crystallization and mutational analysis of PapGII, Waksman and colleagues show that a point mutant W107A reduced the agglutination of human erythrocytes, yet could still expressed as well as wild-type PapGII (*8*). This tryptophan residue forms hydrophobic interactions with the lactose core of the globoceramides and is conserved in all three alleles.

Based on this work, we made the W to A mutation in all three adhesins (PapGI W105A, PapGII W107A, and PapGIII W107A) and assayed their activities compared to the wild-type proteins (Figure 5.4B). As expected, PapGI bound to only to GbO3 (16,225 A.U.) and *i*GbO3 to a minor extent (1,544.5 A.U.). Inhibition with 100 mM lactose reduced the binding by 89% and 75%, respectively against both glycoproteins (Figure 5.4B). On the other hand, PapGI W105A reduced binding to both glycoproteins by 99.6% and 93%, respectively. PapGII WT behaved in a similar manner, but bound both GbO3 and GbO4 the same (21,851 and 22,036 A.U., respectively). Inhibition with lactose again reduced activity by 86% and 88%, respectively (Figure 5.4B). Again, the W107A mutant of PapGII showed less activity when compared to the inhibition experiments (99% and 96% reduction in binding, respectively). In previous work on a different glycoprotein array, PapGIII bound several galactosylated ligands and glycoproteins such as ovalbumin and thyroglobulin (*15*). Surprisingly, PapGIII did not bind to anything on our glycoprotein array, so we could not assay the effect of mutants.

5.2.3 Non-binding mutants of the soluble lectins PA-IL, PA-IIL, and RS-IIL

The calcium-dependent, galactophilic lectin PA-IL derived from the *Pseudomonas aueruginosa* has been widely characterized by our lab and others (*10, 11, 15, 22, 23*). From the initial glycan array screen, PA-IL had a broad preference for terminal α-galactose residues. However, the lectin bound tighter to Galα1-4Gal residues preferentially over Galα1-2/3Gal, and

other α-galactose epitopes with varying internal glycans (*23*). A crystal structure of PA-IL with a bound ligand isoglobotriose (*i*GbO3, Galα1-3Galβ1-4Glc) was solved by Imberty and colleagues highlighting key residues in glycan-binding (*23*) (Figure 5.5A). The binding pocket of PA-IL contains hydrophilic and charged residues that bind the Ca^{2+} ion and the terminal galactose epitope. Residues D100 and N107 bind to both the glycan and the coordinated metal, hence, we mutated both residues to Ala, thereby maintaining structural integrity but losing the electrostatic interactions. Wild-type PA-IL bound well to the galactose-containing glycoproteins GbO3 (64,890 A.U.), α-Gal trisaccharide (64,945A.U.) and Gal (22,520 A.U.), and unexpectedly bound to the GalNAc neoglycoprotein (21,010 A.U.) (Figure 5.5B). Recently, the Gildersleeve group demonstrated the effects of glycan density and multivalency of carbohydrate-presenting carrier proteins (ie BSA) (*24*).

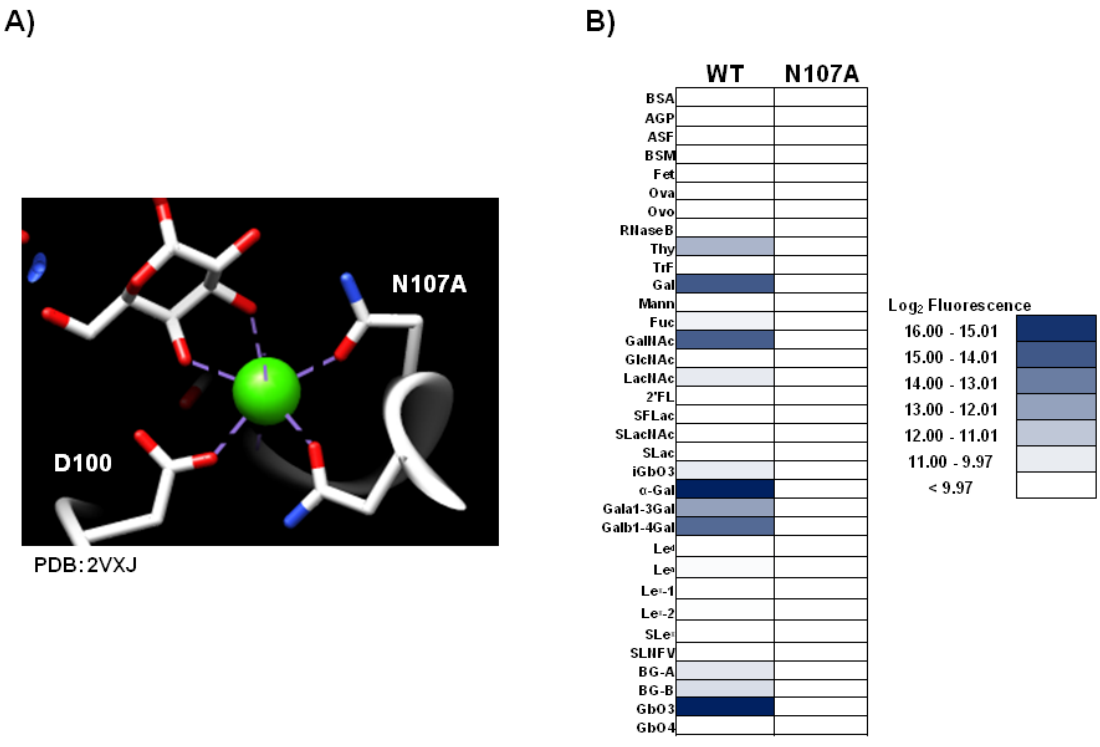

Figure 5.5 Mutational analysis of PA-IL. A) Visualization of the crystal structure of PA-IL with bound GbO3. Close-up of the interactions of N107 and D100 with the terminal α-galactose and Ca^{2+} ion. B) Heat map representation comparing wt PA-IL and PA-IL N107A.

Using our GST-tagged PA-IL, the Gildersleeve group demonstrated that lectin binding to GalNAc-presenting carrier proteins can vary based on the density of the glycans, therefore altering the effective concentration of the presented glycan, resulting in ligand binding (*24*). Although we observe this phenomenon, PA-IL still displays insignificant binding to β-galactose epitopes (LacNAc, Le^a, and Le^x) and others. When PA-IL D100A and PA-IL N107A were hybridized to our array, we observed significantly reduced activity (~98%), with PA-IL N107A displaying the lowest median signal from all glycoproteins (226 A.U. for PA-IL D100A versus 108 A.U. for PA-IL N107A). Henceforth, we used PA-IL N107A as our negative control for PA-IL.

Pseudomonas aeruginosa and *Ralstonia solanacearum* produce lectins, PA-IIL and RS-IIL, respectively, with similar protein sequences and binding specificity (*14*). PA-IIL is mainly a fucose-binding lectin but does show activity against high-mannose epitopes, conversely RS-IIL is a primarily a mannose-binding lectin but does display activity against fucose-containing glycan epitopes (*13*). Crystal structures of both lectins with their appropriate monosaccharides are currently available (*13, 14*). The Imberty group produced three mutants of PA-IIL, one of which, PA-IIL S22A, switches the specificity from a fucose-binding to a mannose-binding lectin (*14*) (Figure 5.6A). RS-IIL WT contains an alanine at position 22, so we theorized that mutating the alanine back to serine, we could alter the specificity of RS-IIL to bind fucose-containing epitopes. Also, both proteins have a conserved aspartate residue critical to binding (PA-IIL D96 and RS-IIL D95). Therefore, we generated PA-IIL S22A, PA-IIL D96A, RS-IIL A22S, and RS-IIL D95A variants and assayed them against our glycoprotein array.

As expected, PA-IIL WT and RS-IIL WT bound very well to their respective glycoproteins (Figure 5.6). PA-IIL WT bound mostly fucosylated epitopes (Lewis and blood

group antigens), and the high-mannose containing RNase B. RS-IIL WT bound high-mannose containing glycoproteins and only mildly bound fucosylated antigens. When PA-IIL D96A was assayed against the panel of glycoproteins displayed significantly reduced activity against all of the positive signals. There was a reduction in binding for RNase B (100% reduced), fucose-BSA (95% reduced), 2'-fucosyllactose (100% reduced), and the Lewis antigens (95 – 98% reduction). RS-IIL D95A did not show any binding activity against any of the glycoproteins tested, exhibiting >98% reduction in binding to all epitopes on the array (Figure 5.6B).

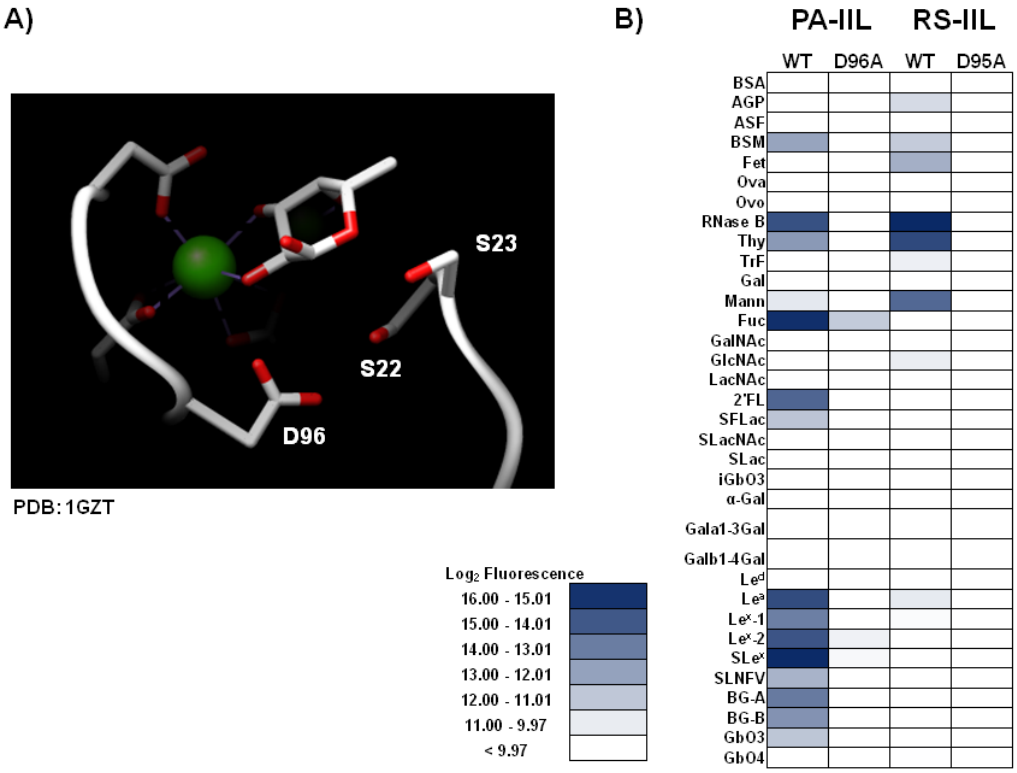

Figure 5.6 Mutational analyses of PA-IIL and RS-IIL. A) Visualization of the crystal structure of PA-IIL with bound fucose. Close-up of the interactions of D96, S22, and S23 with the terminal fucose epitope. B) Heat map representation comparing wt PA-IIL and RS-IIL against the D96/95 analogs.

When profiling the other two mutants, we observed very interesting specificity profiles (Figure 5.7). As described previously, PA-IIL S22A reduced binding to fucose and increased activity for mannose monosaccharides (*14*). When probed against our array, PA-IIL S22A indeed displayed significantly reduced binding to fucosylated epitopes and yet maintained binding activity against RNase B and Mannose-BSA (Figure 5.7). In the case of RS-IIL A22S, we also observed a change of binding compared to the wild-type (Figure 5.7). Although we do observe a reduction in binding to the mannose- containing glycoproteins, we observe increased binding to some fucosylated epitopes (Le^x, Le^a, Sialyl LFP V, and 2'-fucosyllactose). The RS-IIL A22S variant shows binding preference for fucose, which will be examined more in depth by our lab, although it does display reduced binding to the mannosylated ligands. We also performed ELISA assays to determine the specificity of the lectins (Figure 5.8). While probing activity against mannose-BSA and inhibiting with fucose and mannose monosaccharides, we further show that the specificities changed based on the inhibition. For PA-IIL WT, the inhibitions matched the expected outcome as fucose was a much better inhibitor than mannose (3.8 μM versus 61 μM, Figure 5.8A). When testing PA-IIL S22A, we had an observable IC_{50} of 61 μM for both of the monosaccharides (Figure 5.8B). For RS-IIL WT, inhibitions again matched expected outcomes as mannose was a better inhibitor than fucose (1.5 μM versus 976 μM, Figure 5.8C). When testing RS-IIL A22S, we observed an IC_{50} of 1.5 mM when inhibited with mannose, but showed an IC_{50} of 0.39 mM when inhibited with fucose (Figure 5.8D).

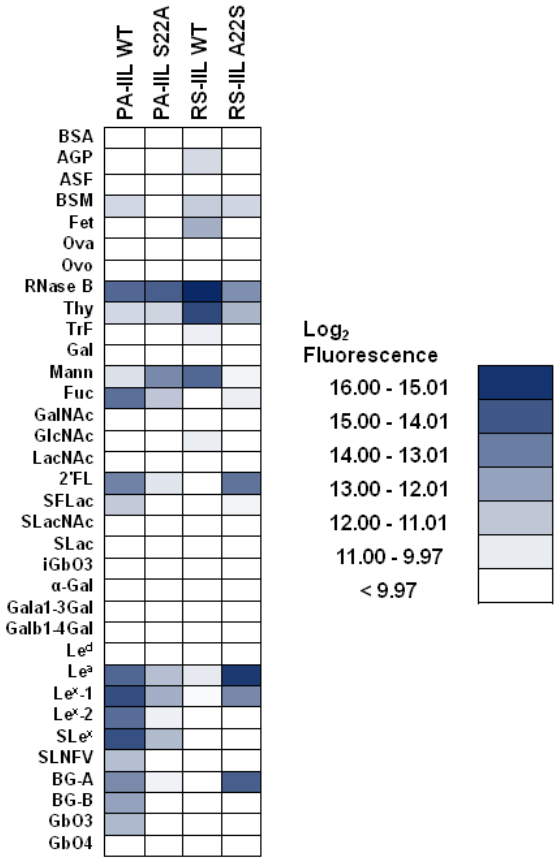

Figure 5.7 Further mutational analyses of PA-IIL and RS-IIL. Heat map representation comparing wt-PA-IIL and RS-IIL against the S22/A22 analogs. For PA-IIL, with the S22A mutant, the specificity changes from solely a fucose-binding lectin to an equal binder of mannose-containing epitopes. For RS-IIL, the A22S mutant reverses specificity from mannose to fucose binding.

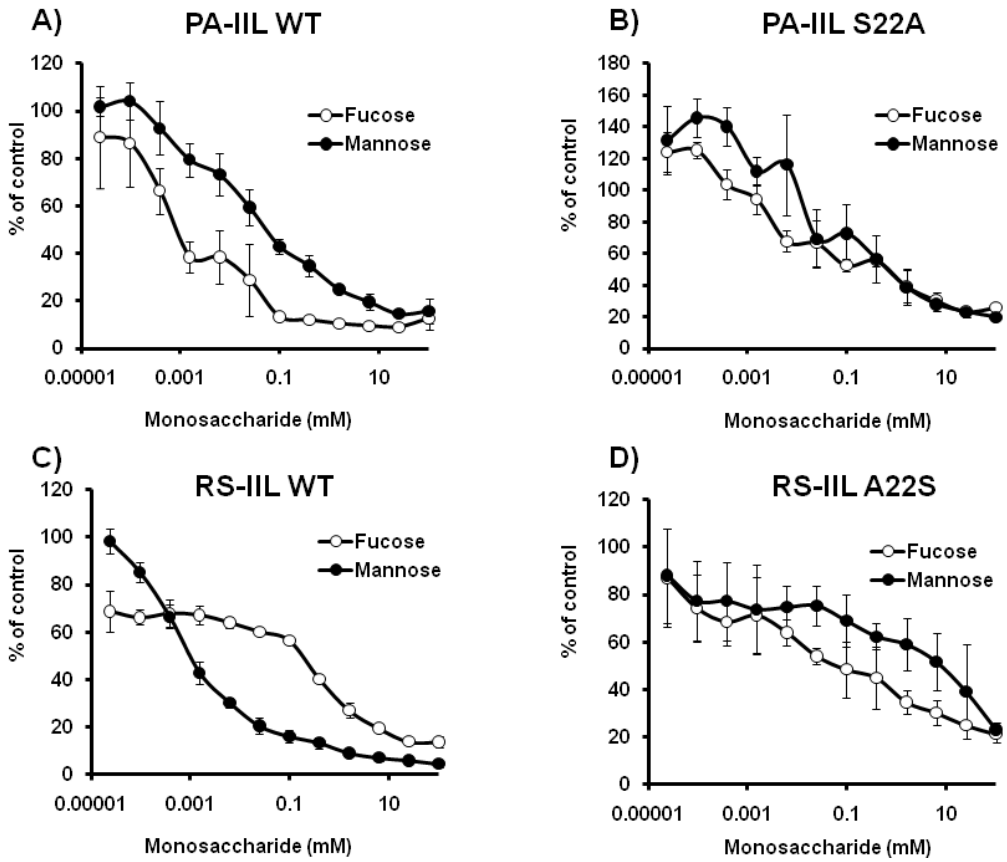

Figure 5.8 ELISA inhibition data for wt PA-IIL, wt RS-IIL, PA-IIL S22A, and RS-IIL A22S against mannose-BSA. A) Inhibition of wt PA-IIL with fucose (open circles) and mannose (closed circles). IC_{50} values were estimated from the inhibition curve. Fucose (3.8 μM) is a better inhibitor of wt PA-IIL than mannose (61 μM).B) Inhibition of PA-IIL S22A with fucose (open circles) and mannose (closed circles). IC_{50} values were estimated from the inhibition curve. Fucose (61 μM) and mannose (61 μM) are very similar inhibitors of glycan-binding. C) Inhibition of wt RS-IIL against fucose (open circles) and mannose (closed circles). IC_{50} values were estimated from the inhibition curve. Mannose (1.5 μM) is a better inhibitor of wt RS-IIL than fucose (976 μM). D) Inhibition of RS-IIL A22S against fucose (open circles) and mannose (closed circles). IC_{50} values were estimated from the inhibition curve. Fucose (0.39 mM) is a better inhibitor of RS-IIL A22S than mannose (1.56 mM). ELISAs was performed in triplicate wells and error bars are standard deviations from the mean signal.

5.2.4 Lectin microarray analysis of lectin variants

Once we were confident with our lectin variants, we wanted to test how the recombinant lectins would bind if they were immobilized into a lectin microarray. In addition to the recombinant lectins and lectin variants which I oriented *in situ*, I also printed a set of 25 plant

lectins in our standard print buffer. I then hybridized the array with a dozen Cy3- or Cy5-labeled glycoproteins, all at 200 nM. We then plotted the data into a heat map, to aid in the visualization of activity (Figure 5.9). As expected, we observed common binding interactions. For the plant lectins, GS-I, HPA, SBA, and SNA bound to galactose epitopes, and DSA and WGA bound to GlcNAc- and GalNAc-containing glycans (Figure 5.9). More interestingly however, UEA-I bound not only the two epitopes with α1-2 fucose linkages, but it also bound to isoglobotriose (iGbO3, Galα1-3Galβ1-4Glc) with relatively high activity. This is odd considering UEA-I is not known to be anything but an α1-2 fucose-binding lectin, although one group has shown that binding to milk oligosaccharides is inhibited by Galα1-3 epitopes (*25*). I further analyzed this binding by inhibiting with fucose, and found that this was dependent of carbohydrate interactions. Looking at the recombinant lectins, specific lectin binding was also observed. For example, GafD bound to GlcNAc, PA-IL bound to the terminal Galα1-4 epitope and also GalNAc, PA-IIL bound fucose- and mannose-containing ligands, and RS-IIL bound to the high-mannose containing glycoproteins OVA, RNase B, and some binding to fucosylated glycans. Unfortunately, at high concentrations of glycoprotein (200 nM) the lectin binding variants displayed activity as well. The variants GafD D88L, PA-IIL D96A, and RS-IIL D95A showed binding to their respective glycans, most likely due to the enhanced activity as a result of *in situ* orientation (Figure 5.9). Although GafD D88L displayed moderate binding in this format, the double mutant GafD D88L + T117A showed significantly reduced binding, arguing that it performs better as a negative control for GafD. Since the mutant showed activity at this concentration of glycoprotein, I also plotted the data with respect to 20 nM glycoprotein concentrations (Figure 5.10). As expected at this lower concentration, the binding mutants showed no activity for their respective glycoproteins (Figure 5.10).

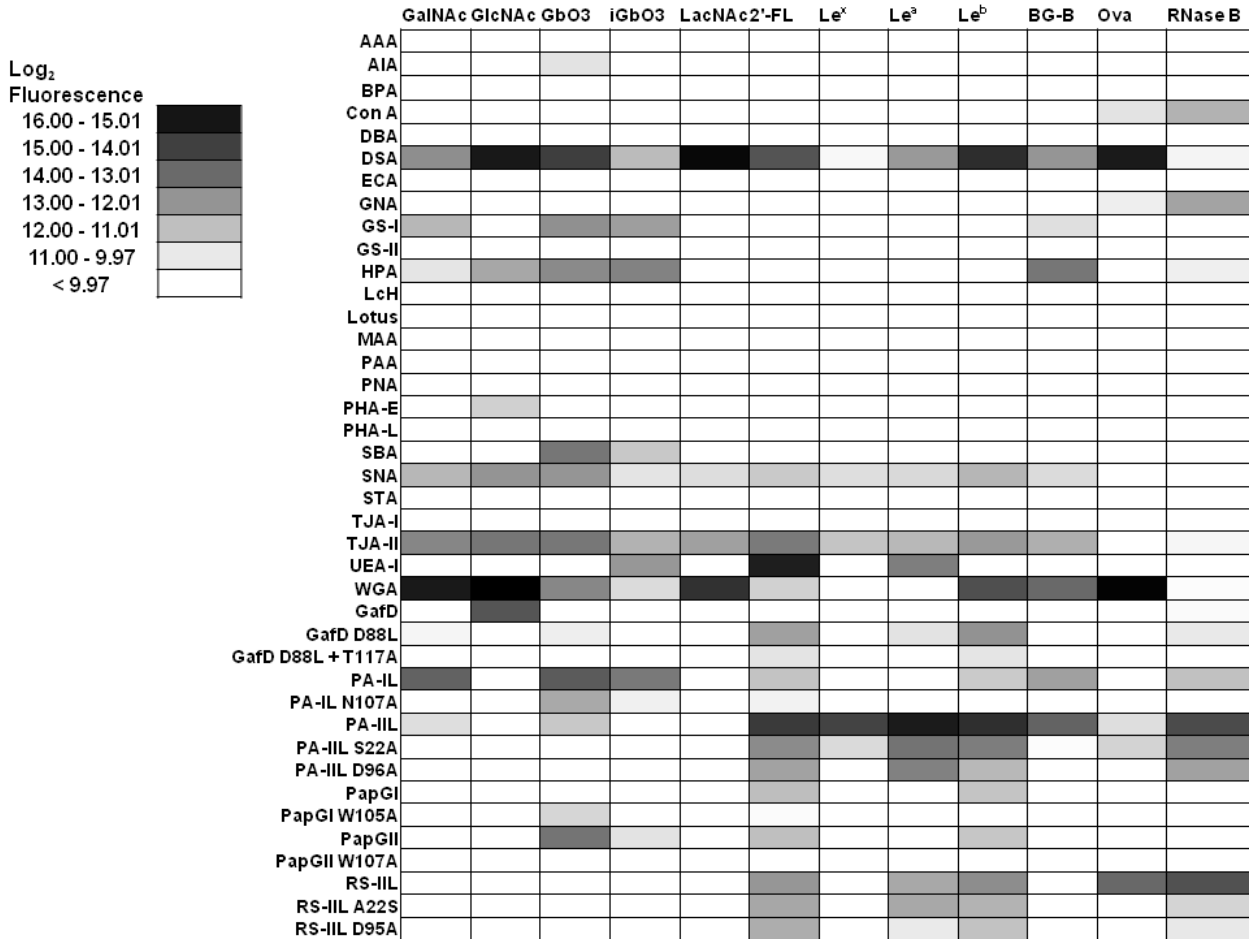

Figure 5.9 Activities of immobilized lectins against 200 nM glycoproteins. Heat map representation of lectin binding activity. Along with the recombinant lectins and variants, a set of 25 plant lectins were also printed.

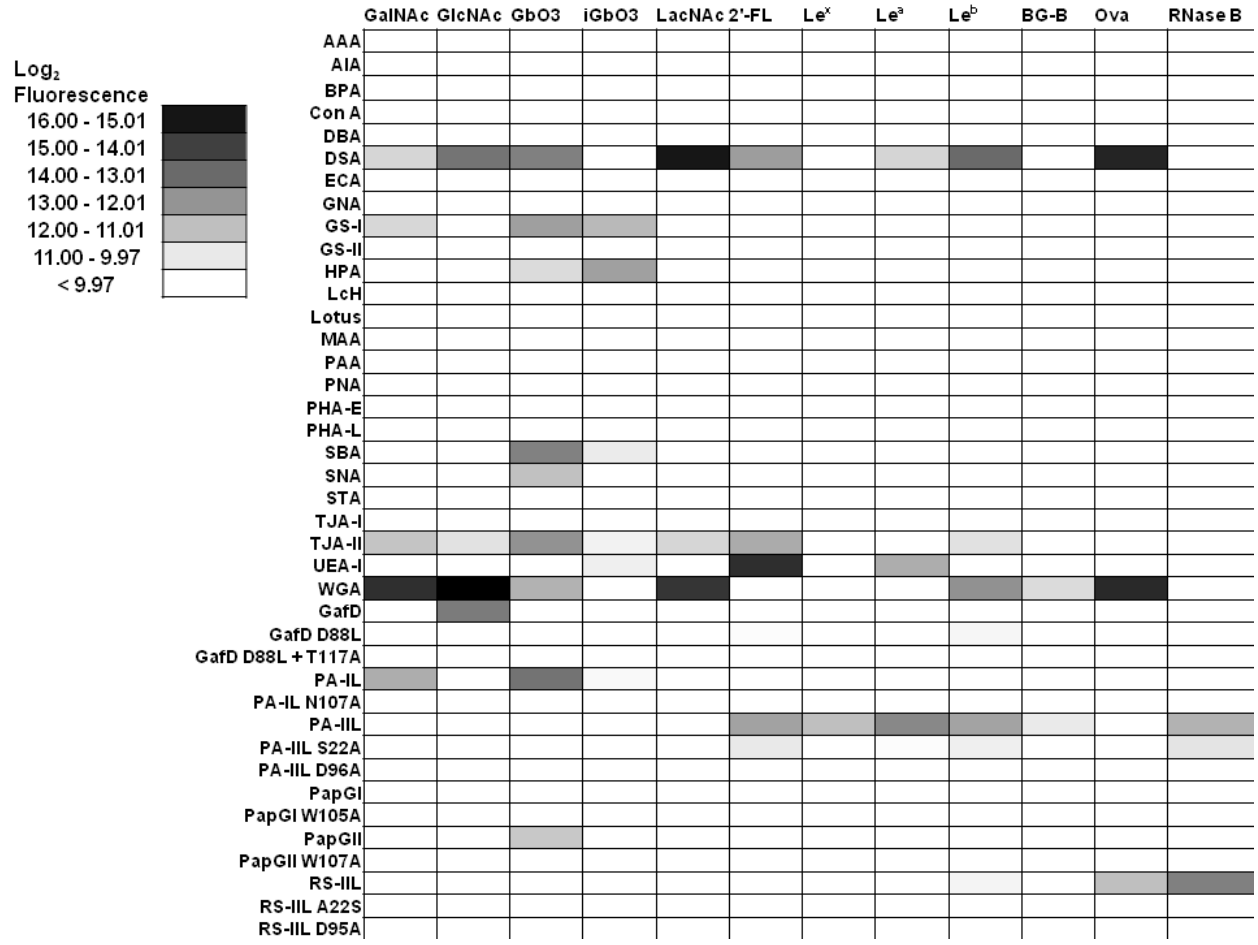

Figure 5.10 Activities of immobilized lectins against 20 nM glycoproteins. Heat map representation of lectin binding activity. Along with the recombinant lectins and variants, a set of 25 plant lectins were also printed.

5.2.5 Current efforts in lectin evolution

The directed evolution of proteins is a common practice in biotechnology to enhance enzyme activity, antibody binding, and specificity (*26*). Evolution is commonly performed by first manipulating the genomic DNA, by either error prone PCR (EP-PCR) or DNA shuffling, followed by selection and/or enrichment of a protein with a specific binding trait (*27*). Despite the need for greater diversity and enhanced specificity of lectins toward a particular glycan epitope, there has been very little work on the development of a technique for the evolution of lectins. One example by the Hirabayashi lab was an attempt to evolve Ricin B-chain lectin from a galactose-binding protein to a sialic acid-binding lectin using EP-PCR and *in vitro* protein synthesis and selection against a sialylated glycoprotein (*28*). Although the group was successful in obtaining a sialic acid-binding lectin, their method was lacking in complete selectivity because the lectin still bound to galactose-containing glycans.

5.2.6 Phage display technology

One of the techniques used to evolve proteins is phage display which requires the expression and display of a given protein onto the surface of a viral particle (*29*). The most prominent carrier virion is the M13 phage, although other phage types are available for manipulation such as the λ- and T4-phage, all of which all *E. coli* specific phage (*30*). The M13 phage is covered by 5 major coat proteins, the predominant coat protein covering the hull of the phage comprises over a thousand copies is the P8 protein (Figure 5.11). Each end of the phage is comprised of a set of two proteins: P3 and P6 on one end and P7 and P9 at the other end (*31, 32*). Through the work of several groups, functional proteins have been produced on each protein type, yet the more predominantly used M13 coat proteins used in phage display

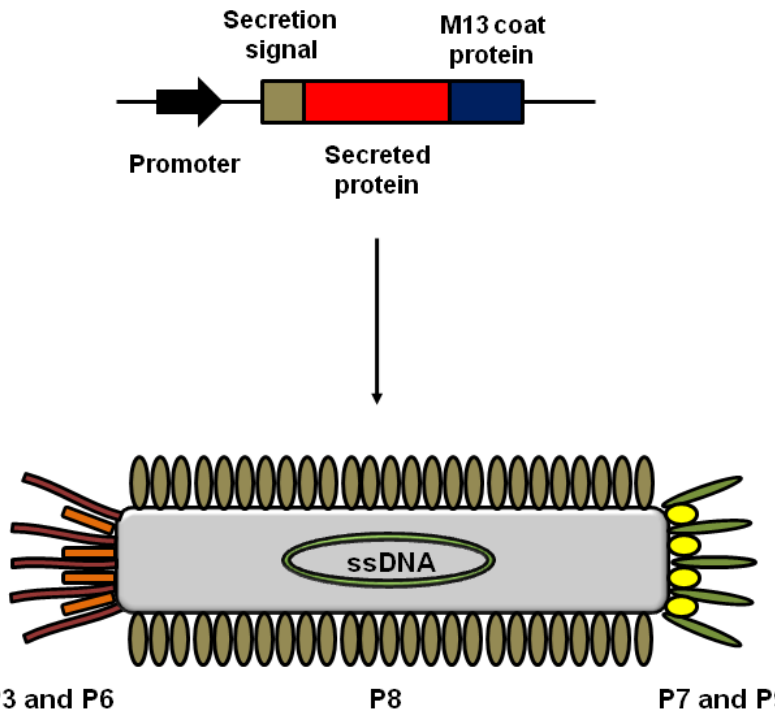

Figure 5.11 Protein expression on the surface of phage. The desired protein is cloned into a phagemid containing a secretion signal, such as pelB, which is fused to an M13 coat protein. When transformed into *E. coli* expressing the F' episome and treated with helper phage, you can obtain whole phage particles with your fusion protein of interest.

are P3 and P8 due to the membrane-bound localization of the protein prior to phage assembly (*33, 34*).

The first examples of phage-displayed polypeptides for the analysis of carbohydrate-binding proteins and glycans were performed using galectin-3 (gal-3) as the displayed protein (*35 - 37*). In one of these studies, Moriki et al. created a library of digested gal-3 DNA and expressed the constructs on the surface of phage, then selected for and identified the smallest fragment needed to bind lactose, a target of gal-3 (*35*). In more recent examples, several groups have focused on polypeptide mimics of carbohydrate antigens to inhibit glycan recognition by lectins (*38*). For example, the Thomsen-Friedenreich (TF)-antigen is a disaccharide unit (Galβ1-3GalNAcα-Ser/Thr) thought to play a role in tumor cell metastasis (*39*). Since the TF-antigen is a known ligand for gal-3, the lectin plays a role in mediating cell adhesion and therefore, tumor

metastasis. Zou et al. demonstrate the selection for a peptide that inhibits gal-3 down to the lower nM range (*37*). In addition, due to the lack of antibodies to the TF-antigen, Heimburg-Molinaro et al. found a peptide that mimicked the glycan, then presented the ligand on a multivalent scaffold. When injected into mice, their TF-mimic induced the production of α-TF-antigen antibodies (*39*).

Given the previous work on lectin evolution, it became clear that using phage-display technology could provide a proper scaffold for the presentation of a lectin. Since there can be issues in expressing plant or eukaryotic proteins, an *E. coli* host seemed the best to produce endogenous recombinant lectins. In the last section of this chapter, I will discuss my efforts toward the directed evolution of GafD (Figure 5.12), a venture we undertook with the supervision and expertise of Professor Jonathan Lai at Albert Einstein College of Medicine.

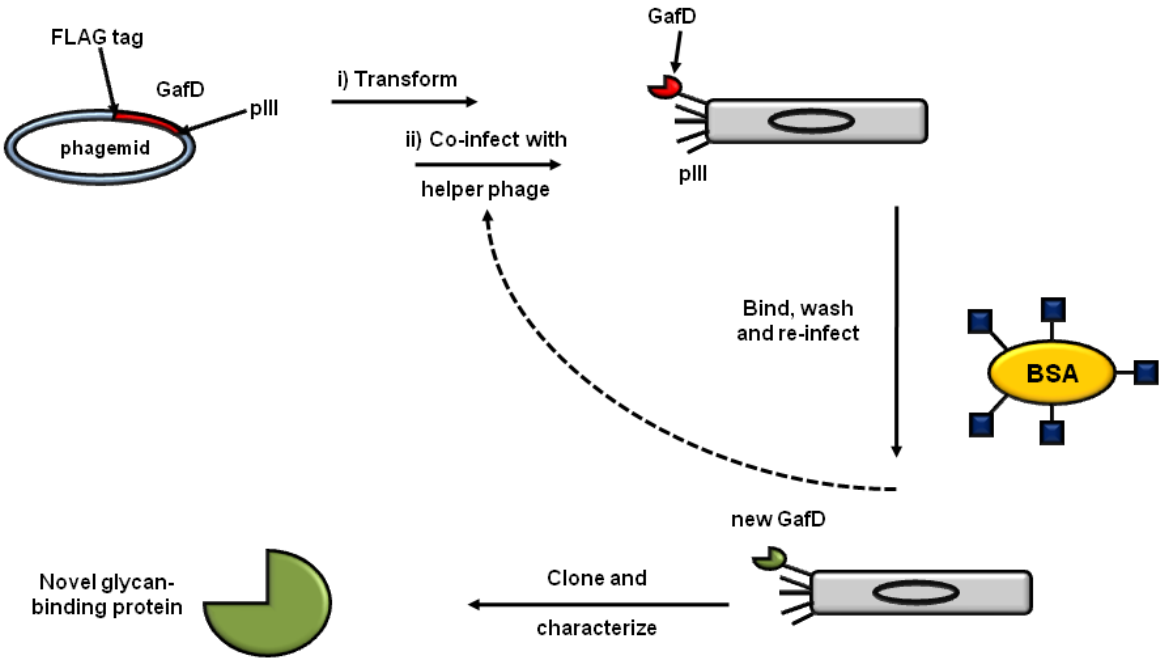

Figure 5.12 Phage display technology for the directed evolution of GafD. GafD would be cloned as a P3 fusion protein, which would then be fused to a phage particle. Creating a library of GafD mutants would allow us to select and enrich a tighter binding 'GafD' to whatever target glycan we decide. Once a candidate is found, the new lectin can be subcloned and purified as a free protein where further studies can be done.

5.2.7 Progress toward the directed evolution of GafD

In the pursuit of exploring binding mutants for the lectins, we became interested in the ability to evolve the recombinant lectins toward specific antigens. Upon arriving to New York University, we were put into contact with Professor Jonathan Lai at Albert Einstein. The focus of the Lai lab is to develop tighter binding single chain variable fragments of antibodies (scFv) to a variety of targets using phage-display technology. Most notably, the group has indentified two inhibitors of the virus-membrane fusion events of HIV-1 and Ebola viruses (*40, 41*). Since the group has expertise in phage-display technology, we sought their help in evolving the GafD lectin. As mentioned before, GafD is a β-GlcNAc-binding adhesin, and we were interested in trying to evolve the lectin to specifically bind β-GalNAc, a C-4 epimer of its native target.

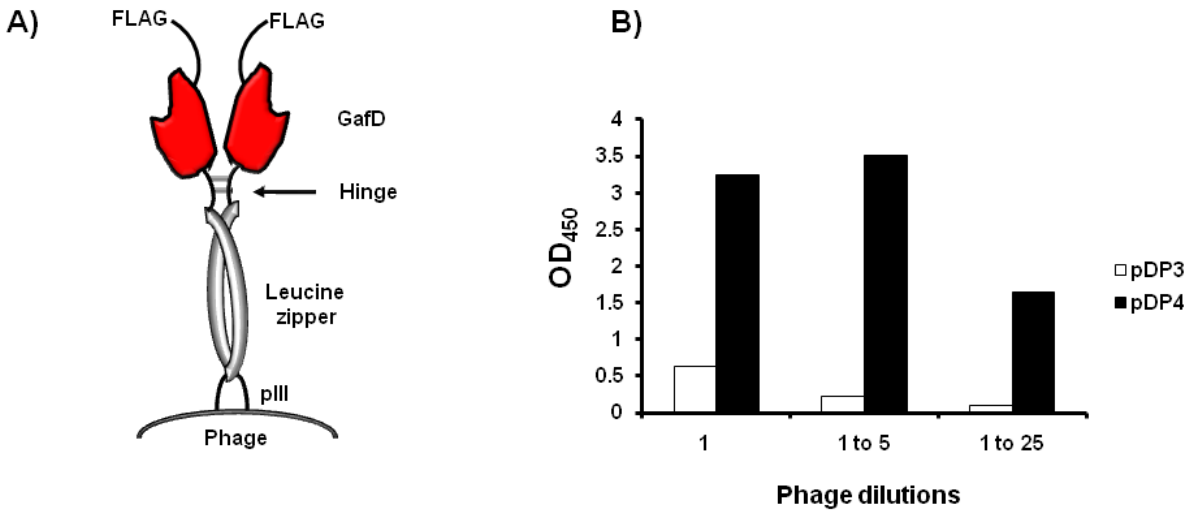

Figure 5.13 Activities of two different phage-displayed methods of GafD. A) Schematic of bivalent displayed GafD. B) Comparison of activity from monovalent GafD (pDP3) and bivalent GafD (pDP4) over several dilutions against 100 ng of β-GlcNAc-HSA.

I cloned GafD into two vectors they had available, a monovalent and bivalent display systems. The monovalent vector (pDP3) was a derivative of a vector developed by the Sullivan lab (*42*), which displays the protein as a fusion with the P3 coat protein. The bivalent vector (pDP4) was created by the Lai lab to enhance protein-ligand avidity for proteins with higher dissociation constants. The fusion protein is supposed to mimic an IgG structure, wherein two phage-displayed proteins would be tethered together *via* a hinge region with disulfide bridges followed by a GCN4, leucine zipper, domain (Figure 5.13A). Upon formation, two fusion proteins would come in contact and bind together creating an "IgG-like" structure. When I tested the activity of the two phage-display systems, the bivalent phage greatly out-performed the monovalent phage due to the increased avidity of the bivalent system (Figure 5.13B). From this point forward, we used bivalent displayed GafD, however, there could be future use of the monvalent GafD which I will discuss later.

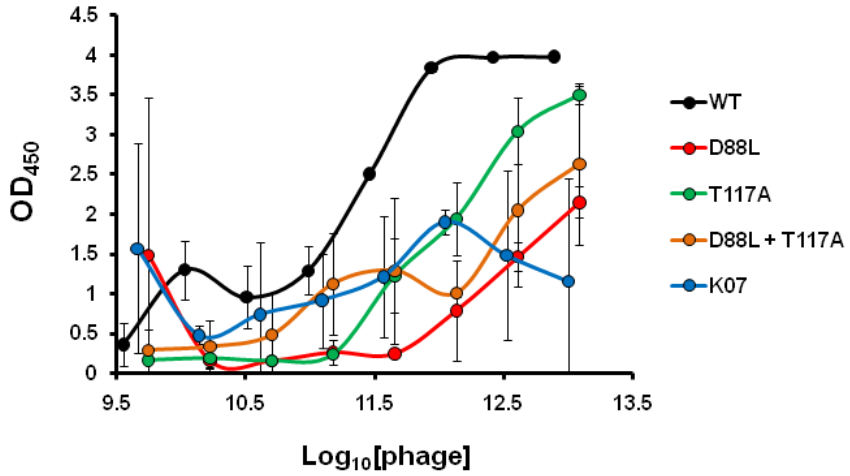

Figure 5.14 Activity of wt GafD against binding variants and a control KO7 helper phage. WT GafD (closed circles) showed the highest activity toward the target, β-GlcNAc-HSA. The mutants D88L (red circles), T117A (green circles), D88L + T117A (orange circles) showed greatly reduced activity compared to the wt. To determine non-specific interactions, we also probed the target with M13KO7 helper phage (blue circles).

To further explore whether GafD could be used in a phage-display system, I made several variants based on my work with the non-binding GafD variants I discovered earlier (Figure 5.3). To obtain mutant DNA, I performed a Kunkel mutagenesis from the corresponding ssDNA (*43*). When the activities were assayed against the same concentration of β-GlcNAc-HSA, wt GafD showed the highest overall activity compared to the three variants, D88L, T117A, and D88L + T117A (Figure 5.14). Helper phage KO7 was used as a second negative control, but the group has found several issues with using the helper phage as a control, so we believe the binding variants function as true negative controls. For further analysis, we compared wt GafD and the T117A variant against BSA, which will act as a negative control during our selections. When tested against BSA, both phage displayed lectins showed minimal binding to BSA, confirming that in a selection, we wouldn't enrich our GafD library against BSA (Figure 5.15A).

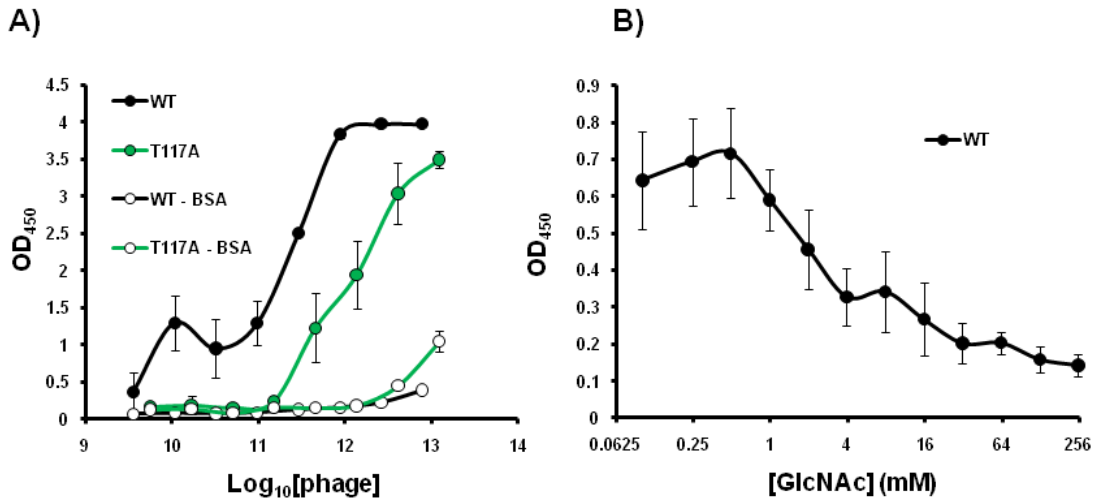

Figure 5.15 Activity of wt GafD against BSA and β-GlcNAc-HSA with inhibiting monosaccharide. A) Activity of wt GafD and T117A variant against β-GlcNAc-HSA (closed circles) and BSA (open circles). B) Inhibition of wt GafD with varying concentrations of *N*-acetylglucosamine (GlcNAc).

We also performed an inhibition experiment using varying amounts of *N*-acetylglucosamine (GlcNAc) and found a concentration dependent signal of phage actvitiy to our target glycan (Figure 5.15B). Since there is some binding of GafD T117A phage to the target protein, we needed to determine if we could enrich wt GafD over the variant phage during a selection. To test this, a made a 1:1000 mixture of wt GafD phage to GafD T117A phage and infected a culture of *E. coli* expressing the F' pilus. After going through five rounds of selection against β-GlcNAc-HSA, we observed complete conversion of the clones present in the output of round 5 from all variant GafD to all wt GafD, suggesting that a selection with varying activities for the target protein can work.

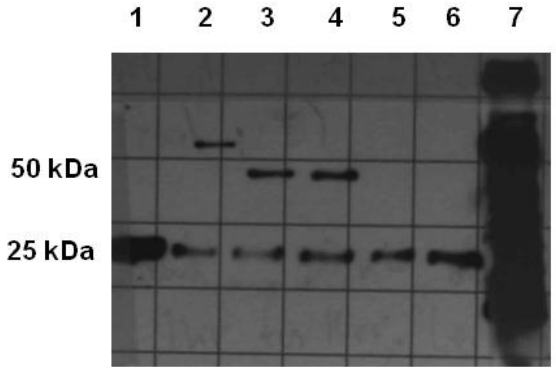

Figure 5.16 Detection of FLAG-tag on phage-displayed GafD. Mono- and bivalent phage ran on a 4 – 12% SDS-PAGE gel, transferred to a Western blot, and probed with 1:3000 α-FLAG-HRP. Lanes: 1) pDP3, 2) KZ52 (a control monovalent sample), 3) pDP4 reduced, 4) pDP4 non-reduced, 5) pLR16 (a control bivalent sample) reduced, 6) pLR16 non-reduced, and 7) FLAG peptide (1 μg).

Since we showed that phage-displayed GafD bound to GlcNAc-BSA, we wanted to determine if we could detect the P3-fusion protein *via* the FLAG tag, which is present in both the mono and bivalent constructs that can then serve as an activity handle in future experiments. Furthermore, if the bivalent phage were run on a gel under non-reducing conditions, we could see fully displayed bivalent protein. The P3 fusion protein has a molecular weight of about 25 kDa and the GafD adhesin has a molecular weight of ~25 kDa. Reduced and non-reduced mono- and bivalent phage were ran on a 4 – 12% SDS-PAGE gel, transferred to a nitrocellulose membrane, and probed with horse-raddish peroxidase conjugated 1:3000 α-FLAG antibody (α-FLAG-HRP), and then visualized using film (Figure 5.16). Unfortunately, for the pDP3, the monovalent phage, I could not see any detectable GafD-P3 fusion protein (Lane 1, Figure 5.16). On the other hand, I did observe this in pDP4 (Lane 3, Figure 5.16). I could not detect any higher order GafD-fusion proteins, meaning I could not detect fully organized bivalent GafD. This may be due to leakage from the sample next to it containing DTT, which can diffuse while preparing

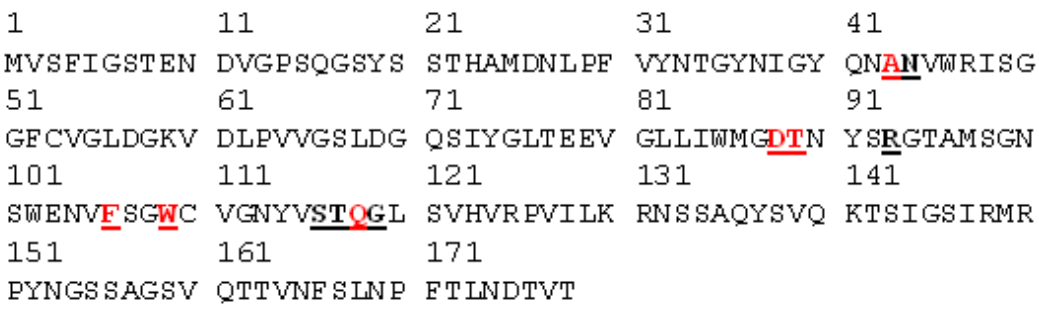

Figure 5.17 Sequence of GafD expressed on the surface of phage. Underlined amino acids were initially mutated into stop codons (TAA), but converted back to their original codon and amino acids in red were randomized (NNS) to make the first GafD library.

the sample (Lane 4, Figure 5.16). Taken together, this data suggests that we have displayed fully functioning GafD on the surface of phage.

Once we knew that GafD has the potential to perform well as a template for library function and selection, we analyzed the protein sequence and found which residues made contact with GlcNAc (*16*). Using Kunkel mutagenesis, we mutated the 11 amino acids chosen to the stop codon TAA. With this template, we can then construct our library. Ideally, when constructing your library all of your primers will anneal to your template DNA, however, this is rarely the case (*43*). By convention, people use a "stop codon template" where your selected amino acids are converted into stop codons, and then your library primers will anneal to this template. Then during rounds of selection, the phage with ssDNA that contain stop codons, a by-product of imperfect annealing conditions, will drop out early in the selection since they should not bind to the target protein. So after making the stop codon template, I then performed the Kunkel mutagenesis with the library primers. Of the 11 amino acids, we decided that we would randomize 6 of the amino acids which are close to the C4 position, to the degenerate 'NNS' codon, which includes 32 different codons of the 20 amino acids, and no stop codons (aside from the amber (UAG) which is translated into glutamate in XL1 Blue cells). This randomization

gives us a theoretical diversity of 1.07 x 10^9 unique codon sets. For the three primers, I obtained a primer incorporation of 42%, which is considered a decent yield. I used a standard annealing method for this Kunkel, but I later learned of an improved Kunkel mutagenesis annealing reaction method (see Materials and methods for both). Upon transformation in pre-treated SS320 cells, a high phage-producing *E. coli* strain developed by the Sidhu lab (*43*), we obtained a 10^9 phage library size.

After I prepped my GafD library phage, I started panning selections against β-GlcNAc-HSA, in the hopes of obtaining a tighter-binding GafD, and β-GalNAc-HSA, the desired target. I initially started out panning against 500 ng of target protein, but when I sequenced my outputs rounds, I observed no convergence of sequence homologies, and more importantly, the stop codon templates that were left over after the Kunkel mutagenesis still populated my output phage to a significant degree (>50% of 40 sequences). Thinking that I have too much target present for the weak binding proteins, I lowered the amount of target protein I used per well to 100 ng. I still went through 5 rounds of selection, yet I noticed something interesting with the selections against β-GalNAc-HSA. I screened all of the output phage and found that R3 gave a significantly higher amount of non-stop codon containing-sequences (~85% positive sequences), whereas R4 and R5 contained higher amounts of stop codon sequences. It was at this point where my research with GafD ended, and passed on to another member of the lab.

5.3 Conclusions

Using the recombinant lectins, we have shown that point mutations of each of the recombinant lectins can produce a binding mutant that does not bind the native antigen. In the case of GafD, we discovered that a double mutant provided the lowest overall signal from all glycans tested. For the PapG adhesins, we found that a mutation of a single amino acid that is

present in all alleles significantly reduces binding. For PA-IL, I made two mutants, yet one, PA-IL N107A, displayed the lowest overall binding to each glycan epitope. In the case of the orthologous lectins PA-IIL and RS-IIL, we show that the same point mutant reduces all binding activity for their respective targets. We also showed that PA-IIL S22A, a mutant produced by the Imberty lab (*14*), shows expected lower fucose binding and maintaining mannose-binding. In contrast, mutating the same residue in RS-IIL back to serine (A22S) resulted in an inverse of binding activity, preferring fucose over mannose residues. With this single point mutant, we achieved the same evolution outcome as the Hirabayashi group presented (*28*), with far less complication. Although we cannot say for certain whether these lectin mutants retain overall protein fold, we believe that substituting the wild-type residues with alanine we could mainting the structure of the binding pocket. However, more work will need to be performed to analyze whether or not we maintin protein folding *via* circular dichroism analysis.

In the pursuit of developing novel binding lectins, we decided to pursue phage-display technology. We decided to modify GafD, the recombinant lectin with the simplest binding motif, in order to bind the β-GalNAc epitope. We enjoyed success in obtaining properly displayed and functional GafD, and created a library of GafD with 10^9 diversity size. We had mild success at our selections, but there are several ways to improve the outcome of our selections. First, we can still tweak the conditions of the selection, such as increasing the amount of target protein coated in the wells to better attract properly folded and bound proteins. Or we can perform a negative selection by first incubating the input phage with BSA, then transferring the supernatant to the target protein. One interesting technique we would try is to incubate the library phage with immobilized α-GafD antibody, which should only pull out properly folded proteins, and would presumably drop out all of the stop codon-containing phage. Another creative method was

developed by the Lai lab to help decrease the amount of stop codon templates present in the later rounds of selection. Instead of the stop codon template, the group used a the rare arginine codon (AGA) which would, during expression, would slow down the production of the "truncated" proteins and allowing full-length proteins to be more favorably expressed (*44*).

5.4 Materials and methods

5.4.1 Cloning, expression, and purification of lectin variants

PapGI genomic DNA was isolated from *Escherichia coli* strain J96 (kind gift from Dr. Rodney A. Welch, University of Madison-Wisconsin) using the DNeasy blood and tissue kit (Qiagen, #69504). The carbohydrate binding domain of PapGI was PCR amplified with the appropriate LIC overhangs: Forward: 5' – GAC GAC GAC AAG ATG GCT GGA TGG CAC AAT GTC – 3'; Reverse: 5' – GAG GAG AAG CCC GGT CAC CCA ACA TTA TCG AAT GA – 3'. PCR reactions were performed with the Hot Start DNA Polymerase (New England Biolabs, #F120S) and reactions were tailored to product specifications. PCR products were treated and ligated into the pET-41 Ek/LIC kit (Novagen, #71017-3) according to the manufacturer's directions. Site-directed mutagenesis was performed using the QuikChange® Lighting Site-Directed Mutagenesis kit (Agilent, #210518-5). Primers (IDT, see Table 5.2) were designed according to product specifications. DNA was then transformed into electrocompetent NovaBlue Gigasingles (Novagen, #71227), grown, and purified DNA was obtained using the Qiaprep Mini Kit (Qiagen, #27106), and sequenced.

Lectin	Sense	Anti-sense
GafD D88L + T117A*	5’ GGA AAT TAT GTA TCA GCG CAG GGA CTG TCT GTT C 3’	5’ G AAC AGA CAG TCC CTG CGC TGA TAC ATA ATT TCC 3’
PapGI W105A	5’ CTT TGT TAA TGG TTA TGA AGC GGA TAC ATG GAC AAA TAA TGG 3’	5’ CCA TTA TTT GTC CAT GTA TCC GCT TCA TAA CCA TTA ACA AAG 3’
PapGII W107A	5’ CTC AAG GGG TAT AAG GCG GAT GAG CGG GCC TTT GAT GCA GGT 3’	5’ ACC TGC ATC AAA GGC CCG CTC ATC CGC CTT ATA CCC CTT GAG 3’
PapGIII W107A	5’ ACC AAG GGA TTT GCA GCG GAA GTC AAC TCA TCT GGA 3’	5’ TCC AGA TGA GTT GAC TTC CGC TGC AAA TCC CTT GGT 3’
PA-IL N107A	5’ GTG CCC GGA ACC TAT GGC GCT AAC TCC GGC TCG TTC AGT GTC 3’	5’ GAC ACT GAA CGA GCC GGA GTT AGC GCC ATA GGT TCC GGG CAC 3’
PA-IIL S22A	5’ GTC ACC GCC TTC GCC AAC GCG TCC GGA ACC CAG ACG 3’	5’ CGT CTG GGT TCC GGA CGC GTT GGC GAA GGC GGT GAC 3’
PA-IIL D96A	5’ GCC CTG GTC GGC TCT GAA GCC GGC ACC GAC AAC GAC TAC AAC 3’	5’ GTT GTA GTC GTT GTC GGT GCC GGC TTC AGA GCC GAC CAG GGC 3’
RS-IIL A22S	5’ GAC GGC ATT TGC AAA TTC AGC GAA CAC CCA GAC 3’	5’ GTC TGG GTG TTC GCT GAA TTT GCA AAT GCC GTC 3’
RS-IIL D95A	5’ GCG ATG GTG GGC TCG GAA GCC GGC ACC GAC AAC GAC 3’	5’ GTC GTT GTC GGT GCC GGC TTC CGA GCC CAC CAT CGC 3’

Table 5.2: Mutagenic primers for recombinant lectins.
* Mutagenesis was performed on GafD D88L DNA.

Purified DNA was transformed into electrocompetent BL21(DE3) cells, and grown on LB-Agar supplemented with kanamycin (30 μg/mL). Colonies were picked for overnight growth in Terrfic Broth (TB, 1.2% tryptone, 2.4% yeast extract, 0.4% glycerol), and were then inoculated into 1 L of TB supplemented with kanamycin (30 μg/mL). Cultures were grown to an OD_{600} 0.7 – 1.0, then induced with 1% lactose and grown for 3 hr at 37°C, shaking at 250 rpm. Cells were pelleted (3000 x g, 15 min) and resuspended in 40 mL of lysis buffer (phosphate buffered saline (PBS, 10 mM sodium phosphate and 15 mM sodium chloride) + 0.2% Triton-

X10) with protease inhibitor mix. Lysozyme (~1 mg/mL) was added and mixed on ice for 30 min. DNase I (New England Biolabs #M0303, ~5 units/mL of lysate) was added and mixed for 10 min on ice. Mixture was then centrifuged at 30,000 x g for 30 min. The supernatant was then loaded onto a 1 mL GSH-sepharose column (GE Healthcare, #17-0756-01), washed with 10 mL of PBS, and eluted with 10 mM GSH in 50 mM Tris, pH 8.0, and collecting 1 mL fractions. Purification was analyzed by SDS-PAGE, and fractions containing sample were pooled and dialyzed against PBS. Aliquots were prepared and snap frozen in liquid N_2 and stored at -80°C until needed.

5.4.2 Glycoprotein microarray protocol

Lyophilized neoglycoproteins (for full name, abbreviations, and suppliers see Table 5.1) were suspended in PBS, and then diluted to 0.5 mg/mL in our glycoprotein print buffer (PBS with 0.005% Tween-20 and 2.5% glycerol) and loaded into a 384-well microplate (Whatman, Piscataway, NJ), and loaded into the SpotBot2 Personal Microarrayer (ArrayIt, Sunnyvale, CA). Printing programs were created with the MMF Spocle Program. Samples were printed with triplicate spots onto Nexterion H slides (Schott North America, Elmsford, NY) with an SMP3 pin (ArrayIt, #SMP3). During the print, the slide chamber maintained humidity at ~50% throughout the print process. After printing, the slides were allowed to warm to room temperature for 2 hr, while maintaining humidity at ~50%. After 2 hr, the slides were then placed in a coplin jar and blocked with 50 mM ethanolamine in 50 mM sodium borate buffer (pH 8.5) for 1 hr, at room temperature with mild shaking. After one hour, the slides were washed with PBS with 0.005% Tween (0.005% PBS-T, 3 x 3 min) and once with PBS. The slides were dried using a slide spinner (Labnet Intl., Edison, NJ), and then fastened in a 24-well hybridization chamber (ArrayIt). The GST-tagged lectins were dissolved in 0.005% PBS-T with 1% BSA, and

were incubated for 1 hr, at roomperature with mild shaking. After 1 hr, the slides were washed with 0.005% PBS-T (5 x 3 min), once with PBS, and then incubated with α-GST rabbit Alexa fluor 488 (10 μg/mL, Invitrogen, #A11131) in 0.005% PBS-T with 1% BSA. The slides were incubated for 1 hr at room temperature with mild shaking. After 1 hr, samples were aspirated and washed with 0.005% PBS-T (5 x 3 min) and once with PBS. The slides were dried as before and loaded into the Genepix 4100A slide scanner (Molecular Devices, Union City, CA). Data was extracted with GenePixPro 5.0 (Molecular Devices) and analyzed and graphed using Microsoft Office Excel 2007.

5.4.3 Lectin mutant ELISA protocol

ELISAs were performed as previously described (*15*). Briefly, 100 μL of Mannose-BSA (Dextra UK) was dissolved in PBS (5 ng/μL) and added to 96-well microtiter plates (Grenier Bio-one, #655061, Monroe, NC). Following incubation (overnight at 4°C), the plates were washed five times with 0.05% PBST (PBS and 0.05% Tween 20). Wells were then blocked with PBS containing 5% BSA for 1 h at RT and again washed five times with PBST. For the inhibition, lectins were diluted (10 ng/μL) in 0.05% PBST containing 1% BSA, 1 mM $CaCl_2$, 1 mM $MgCl_2$, and varying concentrations of the appropriate monosaccharide, and incubated for 30 min. After 30 min, 50 μL of sample were applied to the wells and incubated for another 30 min. In all experiments, wells containing buffer alone were used to measure the background (noise) of the plate. After 30 min, wells were washed five times with PBST and lectin binding was detected with anti-His_6-horseradish peroxidase-conjugated antibody (50 μL, 1 : 625 in 0.05% PBST, 1 h, RT, Novus Biologicals No. NB600-393, Littleton, CO). Wells were then washed 5 times with PBST and HRP activity was detected using *o*-phenylenediamine dihydrochloride (OPD Pierce No. 34 005, Rockford, IL). In brief, OPD solution (100 mL, 0.4 mg/mL in 0.1 M

phosphate/citrate, pH 5.0 containing 0.004% H_2O_2) was added to each well. After 30 min, the enzymatic reaction was stopped by the addition of 50 mL of 2.5 M H_2SO_4. The absorbance at 492 nm was read using a Synergy HT microplate reader (BIO-TEK, Winooski, VT). A reference wavelength of 620 nm was subtracted from these values to account for non-specific absorbance. For data analysis, the lectin binding activity was defined as the signal to noise (S/N): absorbance of samples divided by the average background absorbance values. To determine inhibition of the monosaccharides, percent of control (% control) was calculated: (the absorbance of the sample in the presence of inhibitor/absorbance of uninhibited sample)/100. Microsoft Excel 2007 software was used for statistical analysis, curve fitting, and to generate graphs and tables.

5.4.4 Lectin microarray protocol

Plant and recombinant lectins (Appendix A Table 3) were diluted in PBS and supplemented with 0.5 mg/mL and loaded into a 384-well microplate (Whatman, Piscataway, NJ), and loaded into the SpotBot2 Personal Microarrayer (ArrayIt, Sunnyvale, CA). Printing programs were created with the MMF Spocle Program. Samples were printed with triplicate spots onto Nexterion H slides (Schott North America, Elmsford, NY) with an SMP3 pin (ArrayIt, #SMP3). During the print, the slide chamber maintained humidity at ~50% throughout the print process. After printing, the slides were allowed to warm to room temperature for 2 hr, while maintaining humidity at ~50%. After 2 hr, the slides were then placed in a coplin jar and blocked with 50 mM ethanolamine in 50 mM sodium borate buffer (pH 8.5) for 1 hr, at room temperature with mild shaking. After one hour, the slides were washed with PBS with 0.05% Tween (0.05% PBS-T, 3 x 3 min) and once with PBS. The slides were dried using a slide spinner (Labnet Intl., Edison, NJ), and then fastened in a 24-well hybridization chamber (ArrayIt). Labeled samples were dissolved in 0.005% PBS-T, loaded onto the arrays and incubated for 2 hr

with mild shaking. After 2 hr, the slides were washed with 0.005% PBS-T (5 x 3 min), once with PBS, then dried as before and loaded into the Genepix 4100A slide scanner (Molecular Devices, Union City, CA). Data was extracted with GenePixPro 5.0 (Molecular Devices) and analyzed and graphed using Microsoft Office Excel 2007.

5.4.5 Cloning and Kunkel mutagenesis of GafD for phage display

GafD (see Appendix B for gene) was cloned into the mono- and bi-valent phage display vectors and denoted as pDP3 and pDP4, respectively. For pDP3, I used the following primers with HindIII (5') and SalI (3') restriction sites: Forward: 5' GAC GAC **AAG CTT** GTT TCA TTT ATT GGC AGT ACG GAG 3'; Reverse: 5' GAA GCC **GTC GAC** CTG TGT CAT TCA GCG TAA ATG GAT TCA GGC 3'. For pDP4, I used the following primers with HindIII (5') and SacI (3') restriction sites: Forward: 5' GAC GAC **AAG CTT** GTT TCA TTT ATT GGC AGT ACG GAG 3'; Reverse: 5' GAA GCC **GAG CTC** CTG TGT CAT TCA GCG TAA ATG GAT TCA GGC 3'.

For Kunkel mutagenesis, I first had to prepare deoxyuridine single-stranded DNA (dUssDNA). I transformed pDP4 DNA into electrocompetent CJ236 cells, and plated on LB Agar plates supplemented with chloramphenicol (CAM, 30 μg/mL) and carbenicillin (CARB, 100 μg/mL), and grown overnight at 37°C. The next day, a single colony was picked and grown in 20 mL of 2xYT (in a 250 mL baffle flask) for 1 hr at 37°C. After 1 hr, CARB and CAM was added and incubated for an additional 1 hr. Then 8 μL of M13KO7 (10^{12} concentration) was added, along with 20 μL of uridine (2.5 mg/mL stock) and grown for 1 hr. After 1 hr, kanamycin (KAN, 50 μg/mL) was added and the cultures were grown overnight (~15 hr). After the overnight growth, the culture was transferred to 500 mL 2xYT, supplemented with CARB, CAM, 200 uL of M13KO7, and 50 μL of uridine, then incubated for 1 hr. After 1 hr, KAN was

added, then an additional 50 µL of uridine was added and the cultures were incubated overnight at 37°C (~15 hr). After the overnight incubation, the cultures were pelleted at 7000 rpm for 20 min. The supernatant was transferred to clean centrifuge tubes and then a 5x PEG/NaCl solution was added to make a 1x concentration (4% wt/vol PEG8000, 3% wt/vol NaCl), and incubated one ice for 1 hr. After 1 hr, the solutions were spun at 10,000 rpm for 30 min at 4°C (Note: For these large scale spins, use the 450 mL centrifuge tubes and rotor JA-10). Phage pellets are translucent and were resuspended in PBS. pDP4 dUssDNA was purified following the M13 spin column protocol (Qiagen, #27704).

After quantifying the dUssDNA, Kunkel reactions were ran with 10 µg of total pDP4 dUssDNA. The ratio of mutagenic primer to dUssDNA used was 12:1, and were mixed into a 250 µL reaction solution with the following additives: 25 µL T7 polymerase buffer (New England Biolabs, #M0274), 25 µL T4 DNA ligase buffer (New England Biolabs, #M0202), and added ddi H_2O up to 250 µL total volume. The reaction mixtures were then subjected to a "PCR reaction": 95 °C for 2 min, 42 °C for 3 min, and 25 °C for 5 min. A different cycle optimized later in the Lai lab found that the following cycle resulted in better primer annealing: 90 °C for 5 min, 75 °C for 45 sec, 70 °C for 1 min, 65°C for 1 min, 60 °C for 1 min, 55 °C for 1 min, 52 °C for 3 min, 45°C for 30 sec, 35 °C for 30 sec, 25 °C for 45 sec, and 22 °C for 1.5 min. After the cycle, 10 µL of 25 mM dNTPs (New England Biolabs, #N0447), 3 µL of T7 polymerase, and 2 µL of T4 DNA ligase. The mixtures were incubated at room temperature overnight (~12 hr). The reactions were then purified using the PCR purification kit (Qiagen, #28104) and transformed into XL1 Blue cells. For isolating mutant DNA, single colonies were picked, grown overnight, and mini-prepped. For sequencing libraries, whole LB-agar plants were sent for sequencing and analyzed using BioEdit.

For the stop-codon template, the 5' phosphorylated primers were designed as follows: Primer 1: 5' ACT AAT ACG CCA GAC TTA TTA ATT CTG ATA TCC AAT 3'; Primer 2: 5' GTT TCC ACT CAT CGC GGT ACC TTA GGA ATA ATT TTA TTA CCC CAT CCA TAT AAG 3'; Primer 3: 5' TCT TAC GTG AAC AGA CAG TTA TTA TTA TTA TAC ATA ATT TCC CAC GCA TTA TCC GGA TTA GAC ATT CTC CCA TGA 3'. For the D88L mutant: 5' CCT GGA ATA ATT CGT GAG CCC CAT CCA TAT AAG 3'. For the T117A mutant: 5' G AAC AGA CAG TCC CTG CGC TGA TAC ATA ATT TCC 3'. For the double mutant, both primers were used together since they did not overlap. For the GafD library: Primer 1: 5' ACT AAT ACG CCA GAC ATT SNN ATT CTG ATA TCC AAT 3'; Primer 2: 5' GTT TCC ACT CAT CGC GGT ACC CCT GGA ATA ATT SNN SNN CCC CAT CCA TAT AAG 3'; Primer 3: 5' TCT TAC GTG AAC AGA CAG TCC/SNN SNN CGT TGA TAC ATA ATT TCC CAC GCA SNN TCC GGA SNN GAC ATT CTC CCA TGA 3'. The codons changed to stop codons were either changed to degenerate codons (NNS) or repaired to their original codon. Although I did not have time to perform alanine scanning mutagenesis, I still ordered the following primers: Primer 1: 5' AC TAA TAC GCC AGA CGK YTG CGK YTK SAT ATC CAA TGT TGT A 3'; Primer 2: 5' GT TTC CAC TCA TCG CGG TAC CAS SGG AAT AGK YAG YAK CCC CCA TCC ATA TAA G 3'; Primer 3: 5' TC TTA CGT GAA CAG ACA GAS CTK SAG YGG MAR CAK MAT TTC CCA CGC ACS MTC CGG AAR MGA CAT TCT CCC ATG A 3'.

5.4.6 Creation of GafD library phage

Although I generated enough DNA to finally make a library (~200 μg), I have been informed that one can use ~20 μg of library DNA. First, streak a loop of SS320 cells on LB-agar plates supplemented with tetracycline (TET, 5 μg/mL), and incubate overnight at 37 °C. Inoculate 25 mL of 2xYT containing TET with a single SS320 colony and grow for ~4 hr, until

OD_{600} reaches 0.8. Near the end of the growth, make 10-fold serial dilutions of M13KO7 in PBS (have at least 200 uL to use later). When culture reaches OD_{600} of 0.8, mix 500 µL of SS320 culture and 200 µL of M13KO7 solution into 4 mL of top agar. Then pour the mix on top of pre-warmed, thin, LB-agar TET plates and incubate overnight at 37 °C. Pick a single plaque from the overnight plates, and transfer to 1 mL of 2xYT containing KAN and TET, and incubate for 8 hr at 37 °C. Then, transfer the culture to 20 mL starter culture 2xYT with KAN (in a 250 mL baffled flask) and shake overnight at 37 °C. Then, transfer 5 - 10 mL of the overnight culture to 4 x 500 mL Superbroth supplemented with TET and KAN, and grow until OD_{600} reaches 0.8 (usually ~4 - 5 hr).

The next steps should be performed on ice and in a cold room: Once the cells reach the optimal density, chill the flasks on ice for 5 min with occasional swirling. After 5 min, spin the cells down in 6-500 mL pre-chilled centrifuge tubes (each containing a stir bar) at 5,500 x g, for 10 min at 4 °C. Carefully pour off supernatant, then resuspend pellets in 1 L of 1 mM HEPES buffer (pre-chilled) and combine tubes to 4-500 mL centrifuge tubes. Spin cells down at 5,500 x g, for 10 min, at 4 °C. Carefully pour off supernatant again, then resuspend pellets in 1 L of 1 mM HEPES buffer, and spin down cells again at 5,500 g, for 10 min, at 4 °C. Carefully pour off supernatant, then resuspend pellets in 800 mL of 10% glycerol (pre-chilled), and spin down at 5,500 x g for 10 min at 4 °C. Carefully pour off supernatant again, and resuspend a single pellet in ~150 mL of 10% glycerol. Then transfer suspension to another pellet and resuspend, and repeat until left with one centrifuge tube containing all resuspended pellets. Pellet the cells again, and very careful in pouring off the supernatant, then resuspend the pellet in 2 mL of 10% glycerol (if too thick, add an additional 1 mL). Aliquot in 350 uL amounts and flash freeze and

store at -80 °C. The Lai lab mentions that one can store these aliquots for a couple of days and have not seen any deleterious effects.

I then transformed 1 μg of my library into 350 μL of the prepped SS320 cells and recovered in 5 mL of 2xYT. After recovering for 45 min, I performed serial dilutions of the recovery, staring with 5 μL with 1:10 dilutions, and plated the dilutions onto LB-agar plates with CARB and TET. Incubate overnight at 37 °C and count colonies the next day. I found that 2 colonies had grown on my 10^{-6} plate for my library, and performing the back-calculations, the transformation efficiency of the SS320 cells that I prepped was 10^{9}, suitably competent cells from which to make a library. I then transformed 10 aliquots of the prepared SS320 cells with 1 μg of library DNA each. I recovered with 1 mL of SOC, and then washed each cuvette with an additional 1 mL of SOC, and I transferred all recovered samples to 30 mL of SOC in a 250 mL baffle flask. After are aliquots were pooled, the culture was recovered for 45 min at 37 °C. I then titered the recovery by taking 10 μL of the culture and mixed it with 90 μL of 2YT, then 1:10 dilutions downward. I then plated on LB-agar plates to obtain colonies for sequencing. After recovery, the 50 mL culture was split into 2x500 mL 2xYT cultures in 2 L flasks with CARB and KAN, and grown for ~20 hr at 30 °C. The recovery titer formed 3 colonies at the 10^{-5} dilution, from the back calculations, we have a library the size of 1.5×10^{9}. The recovered cells grew well and then I pelleted the cells as usual, and resuspened the supernatant with 5x PEG/NaCl, and incubated on ice for 1 hr. Flasks were then pelleted at 10,000 rpm, 30 min, 4 °C, then dried as usual, resuspended each pellet in 2 mL of PBS. The library phage was then put into 1 mL aliquots, which can be stored at -80 °C indefinitely with 10% glycerol.

5.4.7 ELISA protocol for phage

β-GlcNAc-BSA (Iso sep) was dissolved into 50 mM sodium borate (pH 8.5, 5 ng/μL) and 100 μL of this solution was added to each well and incubated for 1 hr. After 1 hr, the wells were blocked with 5% BSA in PBS and incubated for an additional 1 hr. The wells were then washed 3x with 0.05 % PBST. Prior to addition of the phage, the samples were diluted into 3% BSA in 0.05% PBST and 100 μL was added to each well and incubated for 1 hr. After 1 hr, the wells were washed 5x with 0.05% PBST, followed by the addition of 100 μL of α-M13-HRP (1:2500 in 1% BSA in 0.05% PBS), and the samples were incubated for 1 hr. The wells were then washed 5x with 0.05% PBST, followed by the addition of 200 μL of TMB solution (Sigma Aldrich, #T0565). After 5 min, the reaction was quenched with 100 μL of 0.5 M H_2SO_4. The plate was then read at 450 nm, and Microsoft Excel 2007 software was used for statistical analysis, curve fitting, and to generate graphs and tables.

5.4.8 Phage selection protocol

At the beginning of the day, inoculate 10 mL of 2xYT, containing TET, with a single colony of XL1 Blue cells, and grow in a 250mL baffled flask at 37 °C until OD_{600} reaches 0.6 – 0.8. For round 1, coat 8 – 10 wells with target protein in 50 mM sodium bicarbonate (pH 8.5, 100 μL total volume per well), and incubate at room temperature for 1 hr. As a control, incubate a single well with 1 μg of BSA to act as a negative control, which will be used to assay library enrichment. After 1 hr, decant the wells and block with 150 μL of 3% BSA in PBS for 1 hr. Then wash wells 3x with 0.05% PBST. After purifying your library phage (R0), mix the phage (100 μL) with 3% BSA in 0.05% PBST (30 μL), and add mixture (100 μL) to each well, and incubate 1 hr. After 1 hr, wash the wells 5x with 0.05% PBST, and after thorough drying, add 100 μL of elution buffer (100 mM glycine, pH 2.0) to each well and incubate for 5 min. After 5 min, pool

elutions into a 1.5 mL tube and add 30 µL of neutralizing buffer (2 M Tris, pH 7.5) per well. Do not mix the negative control well. Once XL1 blue culture has reached an optimal OD_{600}, infect the cells by mixing 5 mL of XL1 blue cells with half of your output phage and incubate for 30 min. After 30 min, add 10 µL of M13KO7 helper phage (10^{12}) and incubate further. Then pour culture into a 250 mL baffle flask containing 25 mL of 2xYT supplememted with CARB and KAN. Repeat selections out to round 5, reducing the number of target wells each round.

5.5 References

1. Buts, L., Bouckaert, J., De Genst, E., Loris, R., Oscarson, S., Lahmann, M., Messens, J., Brosens, E., Wyns, L., and De Greve, H. (2003) The fimbrial adhesin F17-G of enterotoxigenic *Escherichia coli* has an immunoglobulin-like lectin domain that binds *N*-acetylglucosamine. *Mol. Microbiol. 49*, 705 – 715.
2. Vandemaele, F. J., Mugasa, J. P., D. Vandekerchove, D., and Goddeeris, B. M. (2003) Predominance of the *papG*II allele with high sequence homology to that of human isolates among avian pathogenic *Escherichia coli* (APEC). *Vet. Microbiol. 97*, 245 – 257.
3. Stenske K. A., Bemis D. A., Gillespie, B. E., Oliver, S. P., Draughon, F. A., Matteson, K. A., Bartges, J. W. (2009) Prevalence of urovirulence genes *cnf, hlyD, sfa/foc*, and *papGIII* in fecal *Escherichia coli* from healthy dogs and their owners. *Amer. J. Vet. Res.70*, 1401 – 1406.
4. Krogfelt, K. A., Bergmans, H., and Klemm, Per. (1990) Direct evidence that the FimH protein is the mannose-specific adhesin of *Escherichia coli* Type 1 fimbriae. *Infection and Immunity 58*, 1995 – 1998.
5. Smith, C. J., Marron, M. B., Twohig, J. M. G. J., and Smith, S. G. J. (1996) Fimbrial adhesins: similarities and variations in structure and biogenesis. *FEMS Immun. Med. Microbiol. 16*, 127 – 139.
6. De Greve, H., Wyns, L., and Brouckaert, J. (2007) Combining sites of bacterial fimbriae. *Curr. Opin. Struct. Biol. 17*, 506 – 512.
7. Westerlund-Wikstrom, B. and Korhonen, T. K. (2005) Molecular structure of adhesin domains in *Escherichia coli* fimbriae. *Int. J. Med. Microbiol. 295*, 479 – 486.

8. Dodson, K. W., Pinkner, J. S., Rose, T., Magnusson, G., Hultgren, S. J., and Waksman, G. (2001) Structural Basis of the Interaction of the Pyelonephritic *E. coli* Adhesin to Its Human Kidney Receptor. *Cell 105*, 733 – 743.

9. Gilboa-Garber, N. and Sudakevitz, D. (1999) The hemagglutinating activities of *Pseudomonas aeruginosa* lectins PA-IL and PA-IIL exhibit opposite temperature profiles due to different receptor types. *FEMS Immunol. Med. Microbiol. 25*, 365 – 369.

10. Chen, C.-P., Song, S.-C., Gilboa-Garber, N., Chang, K. S. S., and Mu, A. M. (1998) Studies on the binding site of the galactose-specific agglutinin PA-IL from *Pseudomonas aeruginosa*. *Glycobiology 8*, 7 – 16.

11. Sabin, C., Mitchell, E. P., Pokorna, M., Gautier, C., Utille, J.-P., Wimmerova, M. and Imberty, A. (2006) Binding of different monosaccharides by lectin PA-IIL from *Pseudomonas aeruginosa*: Thermodynamic data correlated with X-ray structures. *FEBS Letters 580*, 982 – 987.

12. Winzer, K., Falconer, C., Garber, N. C., Diggle, S. P., Camara, M., and Williams, P. (2000) The *Pseudomonas aeruginosa* lectins PA-IL and PA-IIL are controlled by quorum sensing and by RpoS. *J. Bacteriol. 182*, 6401 – 6411.

13. Sudakevitz, D., Kostlanova, N., Blatman-Jan, G., Mitchell, E. P., Lerrer, B., Wimmerova, M., Katcoff, D. J., Imberty, A., and Gilboa-Garber, N. (2004) A new *Ralstonia solanacearum* high-activity mannose binding lectin RS-IIL structurally resembling *Pseudomonas aeruginosa*. *Mol. Microbiol. 52*, 691 – 700.

14. Adam, J., Pokorna, M., Sabin, C., Mitchell, E. P., Imberty, A., and Wimmerova, M. (2007) Engineering of PA-IIL lectin from *Pseudomonas aeruginosa* – Unraveling the role of the specificity loop for sugar preference. *BMC Struct. Biol. 7*, 36 – 49.

15. Hsu, K.-L., Gildersleeve, J. C., and Mahal, L. K. (2008) A simple strategy for the creation of a recombinant lectin microarray. *Mol. BioSys. 4*, 654 – 662.

16. Merckel, M. C., Tanskanen, J., Edelman, S., Westerlund-Wikstrom, B., Korhonen, T. K., Goldman, A. (2003) The structural basis of receptor-binding *Escherichia coli* associated with diarrhea and septicemia. *J. Mol. Biol. 331*, 897 – 905.

17. Lee, P. S., and Lee, K. H. (2000) Genomic analysis. *Curr. Opin. Biotech. 11*, 171 – 175.

18. Tanskanen, J., Saarela, S., Tankka, S., Rhen, N. M., Korhonen, T. K., Westerlund-Wikstrom, B. (2001) The *gaf* fimbrial gene cluster of *Escherichia coli* expresses a full-size and truncated soluble adhesin protein. *J. Bacteriol. 183*, 512 – 519.

19. Schouppe, D., Rouge, P., Lasanajak, Y., Barre, A., Smith, D. F., Proost, P., Van Damme, E. J. M. (2010) Mutational analysis of the carbohydrate binding activity of the tobacco lectin. *Glycoconj. J. 27*, 613 – 623.

20. Hansson, L., Wallbrandt, P., Andersson, J.-O., Bystrom, M., Backman, A., Carlstein, A., Enquist, K., Lonn, H., Otter, C., and Stromquist, M. (1995) Carbohydrate specificity of the *Escherichia coli* P-pilus papG protein is mediated by its *N*-terminal part. *Biochim. Biophys. Acta 1244*, 377 – 383.

21. Snyder, J. A., Haugen, B. J., Buckles, E. J., Lockatell, C. V., Johnson, D. E., Donnenberg, M. S., Welch, R. A., and Mobley, H. L. T. (2004) Transcriptome of uropathogenic *Escherchia coli* during urinary tract infection. *Infect. and Immun. 72*, 6373 – 6381.

22. Nurisso, A., Blanchard, B., Audfray, A., Rydner, L., Oscarson, S., Varrot, A., and Imberty, A. (2010) Role of water molecules in structure and energetic of *Pseudomonas aeruginosa* lectin I interacting with disaccharides. *J. Biol. Chem. 285*, 20316–20327.

23. Blanchard, B., Nurisso, A., Hollville, E., Tetaud, C., Wiels, J., Pokorna, M., Wimmerova, M., Varrot, A., and Imberty, A. (2008) Structural basis of the preferential binding for globo-series glycosphingolipids displayed by *Pseudomona aeruginosa* lectin I. *J. Mol. Biol. 383*, 837 – 853.

24. Zhang, Y., Li, Q., Rodriguez, L. G., and Gildersleeve, J. C. (2010) An Array-Based Method To Identify Multivalent Inhibitors. *J. Amer. Chem. Soc. 132*, 9653 – 9662.

25. Nakajima, K., Kinoshita, M., Matsushita, N., Urashima, T., Suzuki, M., Suzuki, A., Kakehi, K. (2006) Capillary activity electrophoresis using lectins for the analysis of milk oligosaccharide structure and its application to bovine colostrums oligosaccharides. *Anal. Biochem. 348*, 105 – 114.

26. Bonsor, D. A., and Sundberg, E. J. (2011) Dissecting protein-protein interactions using directed evolution. *Biochemistry 50*, 2394 – 2402.

27. Kittl, R., and Withers, S. G. (2010) New approaches to enzymatic glycoside synthesis through directed evolution. *Carb. Res. 345*, 1272 – 1279.

28. Yabe, R., Suzuki, R., Kuno, A., Fujimoto, Z., Jigami, Y., and Hirabayashi, J. (2007) Tailoring a novel silaic acid-binding lectin from a ricin-B chain-like galactose-binding protein by natural evolution-mimicry. *J. Biochem. 141*, 389 – 399.

29. Bratkovic, T. (2010) Progress in phage display: evolution of the technique and its applications. *Cell. Mol. Life Sci. 67*, 749 – 767.

30. Soderlind, E., Simonsson, A. C., and Borrebaeck, C. A. (1992) Phage display technology in antibody engineering: design of phagemid vectors and *in vitro* maturation systems. *Immunol. Rev. 130*, 109 – 124.

31. Sidhu, S. S. (2001) Engineering M13 for phage display. *Biomol. Engineering 18*, 57 – 63.

32. Sidhu, S. S. (2000) Phage display in pharmaceutical biotechnology. *Curr. Opin. Biotech. 11*, 610 – 616.

33. Sidhu, S. S., and Koide, S. (2007) Phage display for engineering and analyzing protein interaction surfaces. *Curr. Opin. Struct. Biol. 17*, 481 – 487.

34. Weiss, G. A., and Sidhu, S. S. (2000) Design and evolution of artificial M13 coat proteins. *J. Mol. Biol. 300*, 213 – 219.

35. Moriki, T., Kuwabara, I., Liu, F.-T., and Maruyama, I. N. (1999) Protein domain mapping by λ phage display: the minimal lactose-binding domain of galectin-3. *Biochem. Biophys. Res. Comm. 265*, 291 – 296.

36. Yamamoto, M., Kominato, Y., and Yamamoto. (1999) Phage display DNA cloning of protein with carbohydrate activity. *Biochem. Biophys. Res. Comm. 255*, 194 – 199.

37. Zou, J., Glinksky, V. V., Landon, L. A., Matthews, L. and Deutscher, S. L. (2005) Peptides specific to the galectin-3 carbohydrate recognition domain inhibit metastasis-associated cancer cell adhesion. *Carcinogenesis 26*, 309 – 318.

38. Molenaar, T. J. M., Appeldoorn, C. C. M., de Haas, S. A. M., Minchon, I. N., Bonnefoy, A., Hoylaerts, M. F., Pannekoek, H., van Berkel, T. J. C., Kuiper, J., and Biessen, E. A. L. (2002) Specific inhibition of P-selectin-mediated cell adhesion by phage display - derived peptide antagonists. *Blood 100*, 3570 – 3577.

39. Heimburg-Molinaro, J., Almogren, A., Morey, S., Glinksii, O. V., Roy, R., Wilding, G. E., Cheng, R. P., Glinsky, V. V., and Rittenhouse-Olson, K. (2009) Development, characterization, and immunotherapeutic use of peptide mimics of the Thomsen-Friedenreich carbohydrate antigen. *Neoplasia 11*, 780 – 792.

40. Da Silva, G. F., Harrison, J. S., and Lai, J. R. (2010) Contribution of light chain residues to high activity binding in an HIV antibody explored by combinatorial scanning mutagenesis. *Biochemistry 49*, 5464 – 5472.

41. Miller, E. H., Harrison, J. S., Radoshitzky, S. R., Higgins, C. D., Chi, X., Dong, L., Kuhn, J. H., Bavari, S., Lai, J. R., and Chandran, K. (2011) Inhibition of Ebola virus entry by a C-peptide targeted to endosomes. *J. Biol. Chem. 286*, 15854 – 15861.

42. Haidaris, C. G.; Malone, J.; Sherrill, L. A.; Bliss, J. M.; Gaspari, A. A.; Insel, R. A.; Sullivan, M. A. (2002) Recombinant Human Antibody Single Chain Variable Fragments Reactive with Candida Albicans Surface Antigents. *J. Immunol. Methods 257*, 185-202.

43. Tonikian, R., Zhang, Y., Boone, C., and Sidhu, S. S. (2007) Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries. *Nat. Methods 2,* 1368 – 1405.

44. Liu, Y., Stewart, A., Harrison, J. S., Da Silva, G. F., and Lai, J. R. *In submission*. A method for improved selection of synthetic antibodies from phage display libraries produced by Kunkel mutagenesis.

Appendix A: Lectin microarray print lists

Print number	Lectin	[Lectin] µM	[Lectin] µg/mL	Monosaccharide in print buffer
1	GafD	20	1000	GlcNAc
2	GafD	10	500	GlcNAc
3	GafD	5	250	GlcNAc
4	GafD	2.5	125	GlcNAc
5	GafD	1.2	62	GlcNAc
6	GafD	0.6	31	GlcNAc
7	GafD	0.3	16	GlcNAc
8	PA-IL	24	1000	Galactose
9	PA-IL	12	500	Galactose
10	PA-IL	6	250	Galactose
11	PA-IL	3	125	Galactose
12	PA-IL	1.5	62	Galactose
13	PA-IL	0.8	31	Galactose
14	PA-IL	0.4	16	Galactose
15	PA-IIL	24	1000	Fucose
16	PA-IIL	12	500	Fucose
17	PA-IIL	6	250	Fucose
18	PA-IIL	3	125	Fucose
19	PA-IIL	1.5	62	Fucose
20	PA-IIL	0.8	31	Fucose
21	PA-IIL	0.4	16	Fucose
22	PapGII	20	1000	Galactose
23	PapGII	10	500	Galactose
24	PapGII	5	250	Galactose
25	PapGII	2.5	125	Galactose
26	PapGII	1.2	62	Galactose
27	PapGII	0.6	31	Galactose
28	PapGII	0.3	16	Galactose
29	PapGIII	20	1000	Galactose
30	PapGIII	10	500	Galactose
31	PapGIII	5	250	Galactose
32	PapGIII	2.5	125	Galactose
33	PapGIII	1.2	62	Galactose
34	PapGIII	0.6	31	Galactose
35	PapGIII	0.3	16	Galactose
36	RS-IIL	24	1000	Mannose
37	RS-IIL	12	500	Mannose
38	RS-IIL	6	250	Mannose
39	RS-IIL	3	125	Mannose
40	RS-IIL	1.5	62	Mannose

41	RS-IIL	0.8	31	Mannose
42	RS-IIL	0.4	16	Mannose
43	GafD-m	20	1000	GlcNAc
44	GafD-m	10	500	GlcNAc
45	GafD-m	5	250	GlcNAc
46	GafD-m	2.5	125	GlcNAc
47	GafD-m	1.2	62	GlcNAc
48	GafD-m	0.6	31	GlcNAc
49	GafD-m	0.3	16	GlcNAc
50	AAA	14	1000	Fucose
51	Jacalin, AIA	10	500	Galactose
52	BPA	3	500	Galactose
53	ConA	10	500	Mannose
54	DBA	4	500	Galactose
55	DSA	6	500	Lactose
56	ECA	19	1000	Galactose
57	GNA	10	500	Mannose
58	GS-I	4	500	Galactose
59	GS-II	4	500	GlcNAc
60	HPA	6	500	Galactose
61	LcH	20	1000	Mannose
62	Lotus	9	500	Fucose
63	MAA	4	500	Lactose
64	PAA	N/A	500	GlcNAc
65	PNA	5	500	Galactose
66	SBA	4	500	Galactose
67	SNA	3	500	Lactose
68	STA	5	500	GlcNAc
69	UEA-I	8	500	Fucose
70	WGA	28	1000	GlcNAc

Table 1: Three spots were printed for each lectin/concentration and 15 spots/5 lectins were printed per row. Lectins 1 – 70 were printed on the top half of each array, and samples 1 – 49 were printed again on the modified, bottom half of the array.

Number	Row	Lectin	[Lectin] µM	[Lectin] µg/mL	Monosaccharide
1	1	AAA	14	1000	Fucose
2	1	AIA, Jacalin	10	500	Galactose
3	1	BPA	3	500	Galactose
4	1	ConA	10	500	Mannose
5	1	DBA	4	500	Galactose
6	2	DSA	6	500	Lactose
7	2	ECA	19	500	Galactose
8	2	GNA	10	500	Mannose
9	2	GS-I	4	500	Galactose
10	2	GS-II	4	500	GlcNAc
11	3	HPA	6	500	Galactose
12	3	LcH	20	1000	Mannose
13	3	Lotus	9	500	Fucose
14	3	MAA	4	500	Lactose
15	3	PAA	N/A	500	GlcNAc
16	4	PNA	5	500	Galactose
17	4	PHA-E	4	500	Lactose
18	4	PHA-L	5	500	Galactose
19	4	SBA	4	500	Galactose
20	4	SNA	3	500	Lactose
21	5	STA	5	500	GlcNAc
22	5	TJA-I	8	500	Lactose
23	5	TJA-II	8	500	Lactose
24	5	UEA-I	8	500	Fucose
25	5	WGA	28	1000	GlcNAc
26	**6**	**GafD**	**20**	**1000**	**GlcNAc**
27	**6**	**GafD**	**10**	**500**	**GlcNAc**
28	**6**	**GafD**	**5**	**250**	**GlcNAc**
29	**6**	**GafD**	**2.5**	**125**	**GlcNAc**
30	**6**	**GafD**	**1.25**	**62.5**	**GlcNAc**
31	**7**	**PA-IL**	**23**	**1000**	**Galactose**
32	**7**	**PA-IL**	**11**	**500**	**Galactose**
33	**7**	**PA-IL**	**6**	**250**	**Galactose**
34	**7**	**PA-IL**	**3**	**125**	**Galactose**
35	**7**	**PA-IL**	**1.4**	**62.5**	**Galactose**
36	**8**	**PA-IIL**	**23**	**1000**	**Fucose**
37	**8**	**PA-IIL**	**11**	**500**	**Fucose**
38	**8**	**PA-IIL**	**6**	**250**	**Fucose**
39	**8**	**PA-IIL**	**3**	**125**	**Fucose**
40	**8**	**PA-IIL**	**1.4**	**62.5**	**Fucose**
41	**9**	**PapGII**	**19**	**1000**	**Galactose**
42	**9**	**PapGII**	**9**	**500**	**Galactose**
43	**9**	**PapGII**	**4.5**	**250**	**Galactose**

44	**9**	**PapGII**	**2.3**	**125**	**Galactose**
45	**9**	**PapGII**	**1.2**	**62.5**	**Galactose**
46	**10**	**PapGIII**	**19**	**1000**	**Galactose**
47	**10**	**PapGIII**	**9**	**500**	**Galactose**
48	**10**	**PapGIII**	**5**	**250**	**Galactose**
49	**10**	**PapGIII**	**2.5**	**125**	**Galactose**
50	**10**	**PapGIII**	**1.2**	**62.5**	**Galactose**
51	**11**	**RS-IIL**	**23**	**1000**	**Mannose**
52	**11**	**RS-IIL**	**11**	**500**	**Mannose**
53	**11**	**RS-IIL**	**6**	**250**	**Mannose**
54	**11**	**RS-IIL**	**3**	**125**	**Mannose**
55	**11**	**RS-IIL**	**1.4**	**62.5**	**Mannose**

Table 2: Print list for lectin array shown in Figure 4. Three spots were printed for each lectin/concentration and 15 spots/5 lectins were printed per row. Bolded samples were printed in GSH-B (100 mM GSH, 50 mM sodium borate buffer, pH 8.5 with 0.5 mg/mL BSA). All others printed in PB (10 mM sodium phosphate, 15 mM sodium chloride, pH = 7.4 with 0.5 mg/mL BSA)

Number	Lectin	[Lectin] µM	[Lectin] µg/mL	Monosaccharide
1	AAA	14	500	Fucose
2	AIA, Jacalin	10	500	Galactose
3	BPA	3	500	Galactose
4	ConA	10	500	Mannose
5	DBA	4	500	Galactose
6	DSA	6	500	Lactose
7	ECA	19	500	Galactose
8	GNA	10	500	Mannose
9	GS-I	4	500	Galactose
10	GS-II	4	500	GlcNAc
11	HPA	6	500	Galactose
12	LcH	20	500	Mannose
13	Lotus	9	500	Fucose
14	MAA	4	500	Lactose
15	PAA	N/A	500	GlcNAc
16	PNA	5	500	Galactose
17	PHA-E	4	500	Lactose
18	PHA-L	5	500	Galactose
19	SBA	4	500	Galactose
20	SNA	3	500	Lactose
21	STA	5	500	GlcNAc
22	TJA-I	8	500	Lactose
23	TJA-II	8	500	Lactose
24	UEA-I	8	500	Fucose
25	WGA	28	1000	GlcNAc
26	**GafD**	**10**	**500**	**GlcNAc**
27	**GafD D88L**	**10**	**500**	**GlcNAc**
28	**GafD D88L + T117A**	**10**	**500**	**GlcNAc**
29	**PA-IL**	**11**	**500**	**Galactose**
30	**PA-IL N107A**	**11**	**500**	**Galactose**
31	**PA-IIL**	**11**	**500**	**Fucose**
32	**PA-IIL S22A**	**11**	**500**	**Fucose**
33	**PA-IIL D96A**	**11**	**500**	**Fucose**
34	**PapGI**	**9**	**500**	**Galactose**
35	**PapGI W105A**	**9**	**500**	**Galactose**
36	**PapGII**	**9**	**500**	**Galactose**
37	**PapGII W107A**	**9**	**500**	**Galactose**
38	**RS-IIL**	**11**	**500**	**Mannose**
39	**RS-IIL A22S**	**11**	**500**	**Mannose**
40	**RS-IIL D97A**	**11**	**500**	**Mannose**

Table 4: Print list for lectin mutant array. Three spots were printed for each lectin/concentration and 15 spots/5 lectins were printed per row. Bolded samples were printed in GSH-B. All others printed in PB

Appendix B: Genes of synthesized SpG and PpL

SpG Gene Synthesized by Genewiz©:

ATGGAGAAAGAGAAAAAAGTCAAATATTTCCTGCGCAAAAGCGCCTTTGGTCTGGC
ATCAGTTTCAGCAGCATTCCTGGTTGGTAGCACCGTGTTCGCCGTGGATTCACCGAT
TGAGGATACGCCGATCATTCGTAATGGTGGGGAGCTGACAAATCTGCTGGGTAACA
GCGAAACAACACTGGCTCTGCGTAACGAAGAATCAGCAACCGCCGACCTGACAGCG
GCAGCCGTAGCCGATACAGTAGCAGCGGCAGCGGCCGAAAATGCCGGAGCAGCCG
CCTGGGAAGCAGCGGCAGCGGCAGACGCTCTGGCTAAAGCGAAAGCCGACGCCCTG
AAAGAATTCAACAAATATGGCGTGAGCGACTATTATAAAAACCTGATCAATAACGC
TAAAACCGTCGAAGGTGTGAAAGATCTGCAGGCACAGGTGGTAGAAAGTGCCAAAA
AAGCCCGTATTAGTGAAGCGACCGATGGACTGAGTGATTTTCTGAAATCCCAGACAC
CGGCTGAGGACACTGTTAAAAGCATCGAGCTGGCAGAAGCAAAAGTTCTGGCAAAC
CGTGAACTGGACAAATATGGGGTGTCGGACTATCACAAAAACCTGATCAACAATGC
CAAAACGGTTGAGGGGGTTAAAGACCTGCAAGCCCAAGTTGTAGAGAGTGCCAAAA
AAGCCCGTATTAGTGAGGCAACGGACGGCCTGAGTGACTTCCTGAAATCTCAAACA
CCTGCCGAGGATACCGTGAAATCCATTGAACTGGCGGAGGCCAAAGTGCTGGCTAA
CCGTGAGCTGGACAAATATGGTGTAAGCGATTATTATAAAAACCTGATCAATAATGC
CAAAACGGTGGAAGGCGTGAAAGCACTGATCGATGAGATCCTGGCGGCACTGCCTA
AAACAGACACGTATAAACTGATCCTGAACGGCAAAACGCTGAAAGGTGAAACAACC
ACCGAAGCCGTTGACGCCGCCACCGCTGAAAAAGTGTTCAAACAGTATGCCAACGA
TAACGGTGTGGATGGAGAGTGGACCTATGATGACGCCACGAAAACGTTTACAGTGA
CCGAGAAACCGGAAGTGATTGATGCTAGCGAACTGACACCAGCCGTTACCACATAT
AAACTGGTGATTAACGGGAAAACCCTGAAAGGTGAGACTACGACAAAAGCAGTAG
ACGCCGAAACAGCTGAGAAAGCATTCAAACAATATGCTAACGATAATGGGGTGGAC
GGAGTTTGGACGTATGACGACGCCACCAAAACATTCACCGTGACGGAAATGGTTAC
CGAGGTTCCGGGTGACGCTCCTACAGAACCGGAAAAACCAGAAGCGAGCATCCCAC
TGGTTCCTCTGACTCCAGCAACCCCAATTGCCAAAGATGACGCTAAAAAAGATGAC
ACGAAAAAAGAAGATGCTAAAAAACCGGAGGCGAAAAAAGAGGACGCCAAAAAA
GCAGAAACACTGCCGACAACAGGTGAAGGCTCAAATCCGTTCTTTACCGCCGCTGCT
CTGGCTGTTATGGCAGGAGCAGGAGCACTGGCAGTAGCCTCTAAACGTAAAGAGGA
TTGA

PpL gene synthesized by Genewiz ©:
ATGAAAAAAACGGCGATTGCTATTGCTGTGGCTCTGGCAGGATTTGCTACTGTTGCT
CAAGCGGCTGTCGAAAACAAAGAGGAAACCCCGGAAACACCTGAAACCGATTCGG
AAGAAGAAGTGACGATCAAAGCGAACCTGATTTTTGCCAATGGCAGCACCCAAACA
GCGGAATTCAAAGGCACCTTCGAGAAAGCGACCTCTGAAGCATATGCCTATGCCGA
TACGCTGAAAAAAGACAACGGCGAGTATACCGTGGATGTGGCGGATAAAGGTTATA
CCCTGAACATCAAATTTGCCGGTAAAGAGAAAACTCCTGAGGAGCCGAAAGAGGAG
GTTACCATTAAAGCCAATCTGATCTATGCCGACGGAAAAACCCAGACGGCGGAGTT
CAAAGGCACATTCGAAGAAGCAACTGCCGAAGCTTATCGTTATGCTGATGCCCTGA
AAAAAGACAATGGCGAGTATACGGTGGACGTTGCCGACAAAGGCTATACGCTGAAC
ATCAAATTCGCTGGTAAAGAGAAAACCCCAGAAGAACCAAAAGAGGAGGTTACGAT
CAAAGCCAACCTGATCTATGCCGATGGGAAAACACAAACAGCTGAGTTCAAAGGGA
CGTTTGAGGAGGCTACTGCTGAGGCCTATCGCTATGCCGACCTGCTGGCTAAAGAAA
ACGGGAAATATACAGTCGATGTGGCCGACAAAGGTTATACGCTGAACATCAAATTC
GCCGGTAAAGAAAAAACACCGGAGGAGCCTAAAGAAGAAGTCACCATCAAAGCCA
ACCTGATTTATGCCGACGGAAAAACACAAACTGCCGAGTTCAAAGGAACGTTTGCC
GAAGCGACGGCGGAAGCATATCGCTATGCCGATCTGCTGGCCAAAGAGAACGGAAA
ATATACGGCCGACCTGGAAGATGGAGGTTATACAATCAACATTCGTTTCGCCGGTAA
AAAAGTGGACGAGAAACCGGAAGAACCGATGGACACCTATAAACTGATTCTGAACG
GAAAAACGCTGAAAGGCGAGACAACAACCGAAGCCGTGGACGCTGCTACTGCTGA
AAAAGTGTTCAAACAATATGCCAACGACAACGGTGTGGATGGAGAATGGACCTATG
ACGATGCCACCAAAACATTCACAGTGACCGAGAAACCAGAAGTCATTGACGCCTCG
GAACTGACTCCGGCGGTTACAACATATAAACTGGTCATTAACGGAAAAACCCTGAA
AGGCGAGACTACCACAAAAGCGGTAGACGCTGAAACAGCGGAGAAAGCATTCAAA
CAATATGCCAATGATAATGGCGTTGACGGCGTTTGGACATATGACGACGCTACGAA
AACCTTCACGGTGACGGAAATGTAA

References

Adam, J., Pokorna, M., Sabin, C., Mitchell, E. P., Imberty, A., and Wimmerova, M. (2007) Engineering of PA-IIL lectin from *Pseudomonas aeruginosa* – Unraveling the role of the specificity loop for sugar preference. *BMC Struct. Biol. 7*, 36 – 49.

Apweiler, R., Hermjakob, H, and Sharon, N. (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta 1473*, 4 – 8.

Beale, D., and Feinstein A. (1969) Studies on the reduction of a Human 19s immunoglobulin M. *Biochem. J. 112*, 187 – 194.

Berlier, J. E., Rother, A., Buller, G., Bradford, J., Gray, D. R., Filanoski, B. J., Telford, W. G., Yue, S., Liu, J., Cheung, C.-Y., Chang, W., Hirsch, J. D., Beechem, J. M., Haugland, R. P., and Haugland, R. P. (2003) Quantitative comparison of long-wavelength Alexa Fluor dyes to Cy dyes: Fluorescence of the dyes and their bioconjugates. *J. Histochem. Cytochem. 51*, 1699 – 1712.

Blanchard, B., Nurisso, A., Hollville, E., Tetaud, C., Wiels, J., Pokorna, M., Wimmerova, M., Varrot, A., and Imberty, A. (2008) Structural basis of the preferential binding for globo-series glycosphingolipids displayed by *Pseudomona aeruginosa* lectin I. *J. Mol. Biol. 383*, 837 – 853.

Bonsor, D. A., and Sundberg, E. J. (2011) Dissecting protein-protein interactions using directed evolution. *Biochemistry 50*, 2394 – 2402.

Borrebaeck, C. A. K., and Wingren, C. (2009) Design of high-density antibody microarrays of disease proteomics: Key technological issues. *J. Proteomics 72*, 928 – 935.

Bos, P. D., Zhang, X. H., Nadal, C., Shu, W., Gomis, R. R., Nguyen, D. X., Minn, A. J., van de Vijver, M. J., Gerald, W. L., Foekens, J. A., and Massague, J. Genes that mediate breast cancaer metastasis to the brain. *Nature 459*, 1005 – 1009.

Bratkovic, T. (2010) Progress in phage display: evolution of the technique and its applications. *Cell. Mol. Life Sci. 67*, 749 – 767.

Butch, A.W., Macke, K. A., Scott, M. G., Inkster, M., and Nahm, M. H. (1989) Mitogen-induced human IgG subclass expression. II. IgG1 and IgG3 subclasses are preferentially stimulated by a combination of *Staphylococcus aureus* Cowan I and pokeweed mitogen. *Hum. Immunol. 24*, 207 – 218.

Buts, L., Bouckaert, J., De Genst, E., Loris, R., Oscarson, S., Lahmann, M., Messens, J., Brosens, E., Wyns, L., and De Greve, H. (2003) The fimbrial adhesin F17-G of enterotoxigenic *Escherichia coli* has an immunoglobulin-like lectin domain that binds *N*-acetylglucosamine. *Mol. Microbiol. 49*, 705 – 715.

Cathou, R. E., and O'Konski, C. T. (1970) A transient electric birefringence study of the structure of specific IgG antibody. *J. Mol. Biol. 48*, 125 – 131.

Chen, C.-P., Song, S.-C., Gilboa-Garber, N., Chang, K. S. S., and Mu, A. M. (1998) Studies on the binding site of the galactose-specific agglutinin PA-IL from *Pseudomonas aeruginosa*. *Glycobiology 8*, 7 – 16.

Chen, M. L., Adak, A. K., Yeh, N. C., Yang, W. B., Chuang, Y. J., Wong, C. H., Hwang, K. C., Hwu, J. R., Hsieh, S. L., and Lin, C. C. (2008) Fabrication of an oriented Fc-fused lectin microarray through boronate formation. *Angew. Chem. Int. Ed. Engl. 47*, 8627 – 8630.

Cummings, R. (2009) The repertoire of glycan determinants in the human glycome. *Mol. BioSys. 5*, 1087 – 1104.

Damrongchain, D., Yun, K., Kobatake, E., and Aizawa, M. (1997) Self-assembling of glutathione *S*-transferase:calmodulin fusion protein on chemically modified gold surface. *J. Biotech. 55*, 125 – 133.

Da Silva, G. F., Harrison, J. S., and Lai, J. R. (2010) Contribution of light chain residues to high activity binding in an HIV antibody explored by combinatorial scanning mutagenesis. *Biochemistry 49*, 5464 – 5472.

De Greve, H., Wyns, L., and Brouckaert, J. (2007) Combining sites of bacterial fimbriae. *Curr. Opin. Struct. Biol. 17*, 506 – 512.

Deng, Y., Zhu, X.-Y., Kienlen, T., and Guo, A. (2006) Transport at the air/water interface is the reason for rings in protein microarrays. *J. Amer. Chem. Soc. 128*, 2768 – 2769.

Derrick, J. P., and Wigley, D. B. (1992) Crystal structure of a streptococcal protein G domain bound to an Fab fragment. *Nature 359*, 752 – 754.

Dodson, K. W., Pinkner, J. S., Rose, T., Magnusson, G., Hultgren, S. J., and Waksman, G. (2001) Structural Basis of the Interaction of the Pyelonephritic *E. coli* Adhesin to Its Human Kidney Receptor. *Cell 105*, 733 – 743.

Du, W., Ma, X., and Schneider, E. M. (2008) A direct immunoassay assessment of streptavidin- and *N*-hydroxysuccinimide-modified biochips in validation of serological TNFα responses in hemophagocytic lymphohistiocytosis. *J. Biomol. Screening 13*, 515 – 526.

Dube, D. H., and Bertozzi, C. R. (2005) Glycans in cancer and inflammation – potential for therapeutics and diagnosis. *Nat. Rev. Drug Discov. 4*, 477 – 488.

Duncan, M. J., Mann, E. L., Cohen, M. S., Ofek, I., Sharon, N., and Abraham, S. N. (2005) The distinct binding specificities exhibited by enterobacterial type 1 fimbriae are determined by their fimbrial shafts. *J. Biol. Chem. 280*, 37707 – 37716.

Eliasson, M., Olsson, A., Palmcrantz, E., Wiberg, K., Inganas, M., Guss, B., Lindberg, M., and Uhlen, M. (1988) Chimeric IgG-binding receptors engineered from staphylococcal protein A and streptococcal protein G. *J. Biol. Chem. 263*, 4323 – 4327.

El-Sayed, A., Alber, J., Lammler, C., Abdulmawjood, A., Zschock, M., and Castaneda, V. H. (2006) Comparative sequence analysis of *spa* gene of *Staphylococcus aureus* isolated from bovine mastitis: characterization of an unusual *spa* variant. *J. Dairy Res. 73*, 322 – 327.

Fernandes, B., Sagman, U., Auger, M., Demetrio, M., and Dennis, J. W. (1991) β1-6 Branched oligosaccharides as a marker of tumor progression in human breast and colon neoplasia. *Cancer Res. 51*, 718 – 723.

Fernandez, I. C. S., van der Mei, H. C., Lochhead, M. J., Grainger, D. W., and Busscher, H. J. (2007) The inhibition of the adhesion of clinically isolated bacterial strains on multi-component cross-linked poly(ethylene glycol)-based polymer coatings. *Biomaterials 28*, 4105 – 4112.

Fischer, M., Leech, A. P., and Hubbard, R. E. (2011) Comparative assessment of different histidine-tags for immobilization of protein onto surface plasmon resonance sesnsorchips. *Anal. Chem. 83*, 1800 – 1807.

Freeze, H. H. (2006) Genetic defects in the human glycome. *Nat. Rev. Genet. 7*, 537 – 551.

Garber, N., Glick, J., Gilboa-Garber, N., and Heller, A. (1981) Interactions of Pseudomonas aeruginosa lectins with Escherichia coli strains bearing blood group determinants. *J. Gen. Microbiol. 123*, 359 – 363.

Gauvrea, V., Chevallier, P., Vallieres, K., Petitclerc, E., Gaudreault, R. C., and Laroche, G. (2004) Engineering surfaces for bioconjugation: developing strategies and quantifying the extent of the reactions. *Bioconjug. Chem. 15*, 1146 – 1156.

Gilboa-Garber, N. and Sudakevitz, D. (1999) The hemagglutinating activities of *Pseudomonas aeruginosa* lectins PA-IL and PA-IIL exhibit opposite temperature profiles due to different receptor types. *FEMS Immunol. Med. Microbiol. 25*, 365 – 369.

Gu, J., Sato, Y., Kariya, Y., Isaji, T., Taniguchi, N., and Fukuda, T. (2009) A mutual regulation between cell-cell adhesion and N-glycosylation: Implications of the bisecting GlcNAc for biological functions. *J. Proteome Res. 8*, 431 – 435.

Guss, B., Eliasson, M., Olsson, A., Uhlen, M., Frej, A.-K., Jornvall, H., Flock, J.-I., and Lindberg, M. (1986) Structure of the IgG-binding regions of streptococcal protein G. *EMBO J. 5*, 1567 – 1575.

Ha, T. H., Jung, S. O., Lee, J. M., Lee, K. Y., Lee, Y., Park, J. S., and Chung, B. H. (2007) Oriented immobilization of antibodies with GST-fused multiple Fc-specific B-domains on a gold surface. *Anal. Chem. 79*, 546 – 556.

Haab, B. B., Dunham, M. J., and Brown, P. O. (2001) Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Gen. Biol. 2*, 1 – 13.

Haab, B. B. (2003) Methods and applications of antibody microarrays in cancer research. *Proteomics 3*, 2116 – 2122.

Haab, B. B. (2006) Applications of antibody array platforms. *Curr. Opin. Biotech. 17*, 415 – 421.

Haidaris, C. G.; Malone, J.; Sherrill, L. A.; Bliss, J. M.; Gaspari, A. A.; Insel, R. A.; Sullivan, M. A. (2002) Recombinant Human Antibody Single Chain Variable Fragments Reactive with Candida Albicans Surface Antigents. *J. Immunol. Methods 257*, 185-202.

Haltiwanger, R. S., and Lowe, J. B. (2004) Role of glycosylation in development. *Annu. Rev. Biochem. 73*, 491 – 537.

Hansson, L., Wallbrandt, P., Andersson, J.-O., Bystrom, M., Backman, A., Carlstein, A., Enquist, K., Lonn, H., Otter, C., and Stromquist, M. (1995) Carbohydrate specificity of the *Escherichia coli* P-pilus papG protein is mediated by its *N*-terminal part. *Biochim. Biophys. Acta 1244*, 377 – 383.

Harbers, G. M., Emoto, K., Greef, C., Metzger, S. W., Woodward, H. N., Mascali, J. J., Grainger, D. W., and Lochhead, M. J. (2007) Functionalized Poly(ethylene glycol)-based bioassay surface chemistry that facilitates bio-immobilization and inhibits nonspecific protein, bacterial, and mammalian cell adhesion. *Chem. Mater. 19*, 4405 – 4414.

Harlow, E., and Lane, D. eds (1988) Antibodies: A laboratory manual. *Cold Spring Harbor Laboratory, N. Y.*, 617 – 618.

Harper, S., and Speicher, D. W. (2008) Expression and purification of GST fusion proteins. *Curr. Protocols Prot. Sci. Unit 6.6*, 6.6.1 – 6.6.26.

Heimburg-Molinaro, J., Almogren, A., Morey, S., Glinksii, O. V., Roy, R., Wilding, G. E., Cheng, R. P., Glinsky, V. V., and Rittenhouse-Olson, K. (2009) Development, characterization, and immunotherapeutic use of peptide mimics of the Thomsen-Friedenreich carbohydrate antigen. *Neoplasia 11*, 780 – 792.

Heron, B. T., Sateriale, A., Teixeira, J. E., and Huston, C. D. (2011) Evidence for a novel *Entamoeba histolytica* lectin activity that recognizes carbohydrates present on ovalbumin. *Int. J. Parasitol. 41*, 137 – 144.

Hsu, K.-L., Pilobello, K. T., and Mahal, L. K. (2006) Analyzing the dynamic bacterial glycome with a lectin microarray approach. *Nat. Chem. Biol. 2*, 153 – 157.

Hsu, K.-L., and Mahal, L. K. (2006) A lectin microarray approach for the rapid analysis of bacterial glycans. *Nat. Protoc. 1*, 543 – 549.

Hsu, K.-L., Gildersleeve, J. C., and Mahal, L. K. (2008) A simple strategy for the creation of a recombinant lectin microarray. *Mol. BioSys. 4*, 654 – 662.

Hsu, K.-L., Pilobello, K., Krishnamoorthy, L., and Mahal, L. K. (2011) Ratriometric lectin microarray analysis of the mammalian cell surface glycome. *Meth. Mol. Biol. 671*, 117 – 131.

Huang, W., Yang, Q., Umekawa, M., Yamamoto, K., and Wang, L. X. (2010) Anthrobacter endo-β-*N*-acetylglucosaminidase shows transglycosylation activity on complex-type *N*-glycan oxazolines: one-pot conversion of ribonuclease B to sialylated ribonuclease C. *ChemBiochem 11*, 1350 – 1355.

Hucknall, A., Kim, D.-H., Rangarajan, S., Hill, R. T., Reichert, W. M., and Chilkoti, A. (2009) Simple fabrication of antibody microarrays on nonfouling polymer brushes with femtomolar sensitivity for protein analytes in serum and blood. *Adv. Mat. 21*, 1968 – 1971.

Jongsma, M. A., and Litjens, R. H. (2006) Self-assembling protein arrays on DNA chips by auto-labeling fusion proteins with a single DNA address. *Proteomics 6*, 2650 – 2655.

Jung, J.-W., Jung, S.-H., Kim, H.-S., Yuk, J. S., Park, J.-B., Kim, Y.-M., Han, J.-A., Kim, P.-H., and Ha, K.-S. (2006) High-throughput analysis of GST-fusion protein expression and activity-dependent protein interactions on GST-fusion protein arrays with a spectral surface plasmon resonance biosensor. *Proteomics 6*, 1110 – 1120.

Kannan, B., Castelino, K., Chen, F. F., and Majumdar, A. (2006) Lithographic techniques and surface chemistries for the fabrication of PEG-passivated protein microarrays. *Biosensors Bioelectronics 21*, 1960 – 1967.

Kastern, W., Sjobring, U., and Bjorck, L. (1992) Structure of *Peptostreptococcal* protein L and identification of a repeated immunoglobulin light chain-binding domain. *J. Biol. Chem. 267*, 12820 – 12825.

Kato, K., Sato, H., Kim, H. G., and Chung, B. H. (2005) Immbolization of histidine-tagged recombinant proteins onto micropatterned surfaces for cell-based functional assays. *Langmuir 21*, 7071 – 7075.

Kawahashi, Y., Do, N., Takashima, H., Tsuda, C., Oishi, Y., Oyama, R., Yonezawa, M., Miyamoto-Soto, E., and Yanagawa, H. (2003) *In vitro* protein microarrays for detecting protein-protein interactions: Application of a new method for fluorescence labeling of proteins. *Proteomics 3*, 1236 – 1243.

Khan, F., He, M., and Taussig, M. J. (2006) Double-hexahistidine tag with high-activity binding for protein immobilization, purification, and detection on Ni-Nitriloacetic acid surfaces. *Anal. Chem. 78*, 3072 – 3079.

Kittl, R., and Withers, S. G. (2010) New approaches to enzymatic glycoside synthesis through directed evolution. *Carb. Res. 345*, 1272 – 1279.

Kownatzki, E. (1973) Disulfide bonds of human IgM: Differential sensitivity to reductive cleavage. *Scand. J. Immunol. 2*, 433 – 437.

Krishnamoorthy, L., Bess, J. W. Jr., Preston, A. B., Nagashima, K., and Mahal, L. K. (2009) HIV-1 and microvesicles from T cells share a common glycome, arguing for a common origin. *Nat. Chem. Biol. 5*, 244 – 250.

Krishnamoorthy, L., and Mahal, L. K. (2009) Glycomic analysis: an array of technologies. *ACS Chem. Biol. 4*, 715 – 732.

Krogfelt, K. A., Bergmans, H., and Klemm, Per. (1990) Direct evidence that the FimH protein is the mannose-specific adhesin of *Escherichia coli* Type 1 fimbriae. *Infection and Immunity 58*, 1995 – 1998.

Kuno, A., Kato, Y., Matsuda, A., Kaneko, M. K., Ito, H., Amano, K., Chiba, Y., Narimatsu, H., and Hirabayashi, J. (2009) Focused differential glycan analysis with the platform antibody-assisted lectin profiling for glycan-related biomarker verification. *Mol. Cell. Proteomics 8.1*, 99 – 108.

Lee, P. S., and Lee, K. H. (2000) Genomic analysis. *Curr. Opin. Biotech. 11*, 171 – 175.

Liang, C. J., Yamashita, K., and Kobata, A. (1980) Structural study of the carbohydrate moiety of bovine pancreatic ribonuclease B. *J. Biochem. 88*, 51 – 58.

Lesaicherre, M.-L., Lue, R. Y. P., Chen, G. Y. J., Zhu, Q., and Yao, S. Q. (2002) Intein-mediated biotinylation of proteins and its application in a protein microarray. *J. Amer. Chem. Soc. 124*, 8768 – 8769.

Liu, Y., Stewart, A., Harrison, J. S., Da Silva, G. F., and Lai, J. R. *In submission*. A method for improved selection of synthetic antibodies from phage display libraries produced by Kunkel mutagenesis.

MacBeath, G., and Schreiber, S. L. (2000) Printing proteins as microarrays for high-throughput function determination. *Science 289*, 1760 – 1763.

Mahal, L. K. (2008) Glycomics: towards bioinformatic approaches to understanding glycosylation. *Anticancer Agents Med. Chem. 8*, 37 – 51.

Manimala, J. C., Li, Z., Jain, A., Vedbrat, S., and Gildersleeve, J. C. (2005) Carbohydrate array analysis of anti-Tn antibodies and lectins reveals unexpected specificities: Implications for diagnostic and vaccine development. *ChemBiochem 6*, 2229 – 2241.

Masri, M. S., and Friedman, M. (1988) Protein reactions with methyl and ethyl vinyl sulfones. *J. Protein Chem. 7*, 49 – 54.

Mateo, C., Abian, O., Fernandez-Lorente, G., Pedroche, J., Fernandez-Lafuente, R., Guisan, J. M., Tam, A., and Daminati, M. (2002) Epoxy sepabeads: a novel epoxy support for stabilization of industrial enzymes via very intense multipoint covalent attachment. *Biotech. Prog. 18*, 629 – 634.

Matson, R. S., Milton, R. C., Rampal, J. B., Chan, T. S., and Cress, M. C. (2005) Overprint immunoassay using protein A microarrays. *Meth. Mol. Biol. 382*, 273 – 286.

Marth, J. D., and Grewal, P. K. (2008) Mammalian glycosylation immunity. *Nat. Rev. Immunol. 8*, 874 – 887.

Merckel, M. C., Tanskanen, J., Edelman, S., Westerlund-Wikstrom, B., Korhonen, T. K., Goldman, A. (2003) The structural basis of receptor-binding *Escherichia coli* associated with diarrhea and septicemia. *J. Mol. Biol. 331*, 897 – 905.

Miller, E. H., Harrison, J. S., Radoshitzky, S. R., Higgins, C. D., Chi, X., Dong, L., Kuhn, J. H., Bavari, S., Lai, J. R., and Chandran, K. (2011) Inhibition of Ebola virus entry by a C-peptide targeted to endosomes. *J. Biol. Chem. 286*, 15854 – 15861.

Molenaar, T. J. M., Appeldoorn, C. C. M., de Haas, S. A. M., Minchon, I. N., Bonnefoy, A., Hoylaerts, M. F., Pannekoek, H., van Berkel, T. J. C., Kuiper, J., and Biessen, E. A. L. (2002) Specific inhibition of P-selectin-mediated cell adhesion by phage display -derived peptide antagonists. *Blood 100*, 3570 – 3577.

Moriki, T., Kuwabara, I., Liu, F.-T., and Maruyama, I. N. (1999) Protein domain mapping by λ phage display: the minimal lactose-binding domain of galectin-3. *Biochem. Biophys. Res. Comm. 265*, 291 – 296.

Nakajima, K., Kinoshita, M., Matsushita, N., Urashima, T., Suzuki, M., Suzuki, A., Kakehi, K. (2006) Capillary activity electrophoresis using lectins for the analysis of milk

oligosaccharide structure and its application to bovine colostrums oligosaccharides. *Anal. Biochem. 348*, 105 – 114.

Nishihara, T., Maeda, H., Okamoto, K.-I., Oshida, T., Mizoguchi, T., and Terada, T. (1991) Inactivation of human placenta Glutathione-*S*-Transferase by SH/SS exchange reaction with biological sulfides. *Biochim. Biophys. Res. Comm. 174*, 580 – 585.

Nurisso, A., Blanchard, B., Audfray, A., Rydner, L., Oscarson, S., Varrot, A., and Imberty, A. (2010) Role of water molecules in structure and energetic of *Pseudomonas aeruginosa* lectin I interacting with disaccharides. *J. Biol. Chem. 285*, 20316–20327.

Ohtsubo, K., and Marth, J. D. (2006) Glycosylation in cellular mechanisms of health and disease. *Cell 126*, 855 – 867.

Patwa, T. H., Wang, Y., Miller, F. R., Goodison, S., Pennathur, S., Barder, T. J., and Lubman, D. M. (2008) A novel phosphoprotein analysis scheme for assessing changes in premalignant and malignant breast cell lines using 2D liquid separations, protein microarrays and tandem mass spectrometry. *Proteomics Clin. Appl. 3*, 51 – 66.

Pavlickova, P., Knappik, A., Kambhampati, D., Ortigo, F., and Hug, H. (2003) Microarray of recombinant antibodies using a streptavidin sensor surface self-assembled onto a gold layer. *BioTechniques 34*, 124 – 130.

Peluso, P., Wilson, D. S., Do, D., Tran, H., Venkatasubbaiah, M., Quincy, D., Heidecker, B., Poindexter, K., Tolani, N., Phelan, M., Witte, K., Jung, L. S., Wagner, P., and Nock, S. (2003) Optimizing antibody immobilization strategies for the construction of protein microarrays. *Anal. Biochem. 312*, 113 – 124.

Pilobello, K. T., Krishnamoorthy, L., Slawek, D. and Mahal, L. K. (2005) Development of a lectin microarray for the rapid analysis of protein glycopatterns. *ChemBiochem 6*, 985 – 989.

Pilobello, K. T., and Mahal, L. K. (2007) Deciphering the glycocode: the complexity and analytical challenge of glycomics. *Curr. Opin. Chem. Biol. 11*, 300 – 305.

Pilobello, K. T., and Mahal, L. K. (2007) Lectin microarrays for glycoprotein analysis. *Meth. Mol. Biol. 385*, 193 – 203.

Pilobello, K. T., Slawek, D. E., and Mahal, L. K. (2007) A ratiometric lectin microarray approach to analysis of the dynamic mammalian glycome. *Proc. Natl. Acad. Sci. 104*, 11534 – 11539.

Plaut, A. G., and Tomasi, T. B. Jr. (1970) Immunoglobulin M: pentameric Fcmu fragments released by trypsin at higher temperatures. *Proc. Natl. Acad. Sci. 65*, 318 – 322.

Propheter, D. C., Hsu, K.-L., and Mahal, L. K. (2010) Fabrication of an oriented lectin microarray. *ChemBiochem 11*, 1203 – 1207.

Propheter, D. C., Hsu, K.-L., and Mahal, L. K. (2011) Recombinant lectin microarrays for glycomic analysis. *Meth. Mol. Biol. 723*, 67 – 77.

Propheter, D. C., and Mahal, L. K. (2011) Orientation of GST-tagged lectins *via in situ* surface modification to create an expanded lectin microarray for glycomic analysis. *Mol. BioSys. 7*, 2114 – 2117.

Rakus, J. F., and Mahal, L. K. (2011) New technologies for glycomic analysis: Toward a systematic understanding of the glycome. *Annu. Rev. Anal. Chem. 4*, 367 – 392.

Reichel, A., Schaible, D., Al Furoukh, N., Cohen, M., Schreiber, G., and Piehler, J. (2007) Noncovalent, site-specific biotinylation of histidine-tagged proteins. *Anal. Chem. 79*, 8590 – 8600.

Rizzi, S. C., and Hubbell, J. A. (2005) Recombinant protein-co-PEG networks as cell-adhesive and proteolytically degradable hydrogel matrixes. Part I: Development and physicochemical characteristics. *Biomacromolecules 6*, 1226 – 1238.

Ro, H.-S., Jung, S. O., Kho, B. H., Hong, H. P., Lee, J. S., Shin, Y.-B., Kim, M. G., and Chung, B. H. (2005) Surface plasmon resonance imaging-based protein array chip system for monitoring a hexahistidine-tagged protein during expression an purification. *Appl. Environ. Microbiol. 71*, 1089 – 1092.

Rusmini, F., Zhong, Z., and Feijen, J. (2007) Protein immobilization strategies for protein biochips. *Biomacromolecules 8*, 1775 – 1789.

Sabin, C., Mitchell, E. P., Pokorna, M., Gautier, C., Utille, J.-P., Wimmerova, M. and Imberty, A. (2006) Binding of different monosaccharides by lectin PA-IIL from *Pseudomonas aeruginosa*: Thermodynamic data correlated with X-ray structures. *FEBS Letters 580*, 982 – 987.

Schott North America Inc. (2009) Protocol: Nexterion® slide H protein application. http://www.us.schott.com/nexterion.

Schouppe, D., Rouge, P., Lasanajak, Y., Barre, A., Smith, D. F., Proost, P., Van Damme, E. J. M. (2010) Mutational analysis of the carbohydrate binding activity of the tobacco lectin. *Glycoconj. J. 27*, 613 – 623.

Sharon, N., and Ofek, I. (2000) Safe as mother's milk: carbohydrates as future anti-adhesion drugs for bacterial diseases. *Glycoconj. J. 17*, 659 – 664.

Sharon, N. (2006) Carbohydrates as future anti-adhesion drugs for infectious diseases. *Biochim. Biophys. Acta 1760*, 527 – 537.

Sidhu, S. S. (2000) Phage display in pharmaceutical biotechnology. *Curr. Opin. Biotech. 11*, 610 – 616.

Sidhu, S. S. (2001) Engineering M13 for phage display. *Biomol. Engineering 18*, 57 – 63.

Sidhu, S. S., and Koide, S. (2007) Phage display for engineering and analyzing protein interaction surfaces. *Curr. Opin. Struct. Biol. 17*, 481 – 487.

Smith, C. J., Marron, M. B., Twohig, J. M. G. J., and Smith, S. G. J. (1996) Fimbrial adhesins: similarities and variations in structure and biogenesis. *FEMS Immun. Med. Microbiol. 16*, 127 – 139.

Snyder, J. A., Haugen, B. J., Buckles, E. J., Lockatell, C. V., Johnson, D. E., Donnenberg, M. S., Welch, R. A., and Mobley, H. L. T. (2004) Transcriptome of uropathogenic *Escherchia coli* during urinary tract infection. *Infect. and Immun. 72*, 6373 – 6381.

Soderlind, E., Simonsson, A. C., and Borrebaeck, C. A. (1992) Phage display technology in antibody engineering: design of phagemid vectors and *in vitro* maturation systems. *Immunol. Rev. 130*, 109 – 124.

Stenske K. A., Bemis D. A., Gillespie, B. E., Oliver, S. P., Draughon, F. A., Matteson, K. A., Bartges, J. W. (2009) Prevalence of urovirulence genes *cnf, hlyD, sfa/foc*, and *papGIII* in fecal *Escherichia coli* from healthy dogs and their owners. *Amer. J. Vet. Res.70*, 1401 – 1406.

Sudakevitz, D., Kostlanova, N., Blatman-Jan, G., Mitchell, E. P., Lerrer, B., Wimmerova, M., Katcoff, D. J., Imberty, A., and Gilboa-Garber, N. (2004) A new *Ralstonia solanacearum* high-activity mannose binding lectin RS-IIL structurally resembling *Pseudomonas aeruginosa*. *Mol. Microbiol. 52*, 691 – 700.

Taeusch, H. W., Ballard, R. A., Gleason, C. A., and Avery, M. E. (2005) Avery's diseases of the newborn. *Elsevier Health Sciences, 8th Ed. Ch. 35*, 454.

Tanskanen, J., Saarela, S., Tankka, S., Rhen, N. M., Korhonen, T. K., Westerlund-Wikstrom, B. (2001) The *gaf* fimbrial gene cluster of *Escherichia coli* expresses a full-size and truncated soluble adhesin protein. *J. Bacteriol. 183*, 512 – 519.

Tateno, H., Uchiyama, N., Kuno, A., Togayachi, A., Sato, T., Narimatsu, H., and Hirabayashi, J. (2007) A novel strategy for mammalian cell surface glycome profiling using lectin microarray. *Glycobiology 17*, 1138 – 1146.

Tateno, H., Toyota, M., Saito, S., Onuma, Y., Ito, Y., Hiemori, K., Fukumura, M., Matsushima, A., Nakanishi, M., Ohnuma, K., Akutsu, H., Umezawa, A., Horimoto, K., Hirabayashi, J., Asashima, M. (2011) Glycome diagnosis of human induced pluripotent stem cells using lectin microarray. *J. Biol. Chem. 286*, 20345 – 20353.

Tomizaki, K., Usui, K., and Mihara, H. (2010) Protein-protein interactions and selection: array-based techniques for screening disease-associated biomarkers in predictive/early diagnosis. *FEBS J. 277*, 1996 – 2005.

Tonikian, R., Zhang, Y., Boone, C., and Sidhu, S. S. (2007) Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries. *Nat. Methods 2,* 1368 – 1405.

Toyoda, M., Yamazaki-Inoue, M., Itakura, Y., Kuno, A., Ogawa, T., Yamada, M., Akutsu, H., Takahashi, Y., Kanzai, S., Narimatsu, H., Hirabayashi, J., and Umezawa, A. (2011) Lectin microarray analysis of pluripotent and multipotent stem cells. *Genes Cells 16*, 1 – 11.

Uchiyama, N., Kuno, A., Tateno, H., Kubo, Y., Mizuno, M., Noguchi, M., and Hirabayashi, J. (2008) Optimization of evanescent-field fluorescence-assisted lectin microarray for high-

sensitivity detection of monovalent oligosaccharides and glycoproteins, *Proteomics 8*, 3042 – 3050.

Valiokas, R., Klenkar, G., Tinazli, A., Tampe, R., Liedberg, B., and Piehler, J. (2006) Differential protein assembly on micropatterned surfaces with tailored molecular and surface multivalency. *ChemBiochem 7*, 1325 – 1329.

Vandemaele, F. J., Mugasa, J. P., D. Vandekerchove, D., and Goddeeris, B. M. (2003) Predominance of the *papG*II allele with high sequence homology to that of human isolates among avian pathogenic *Escherichia coli* (APEC). *Vet. Microbiol. 97*, 245 – 257.

Vareiro, M. M. L. M., Liu, J., Knoll, W., Zak, K., Williams, D., and Jenkins, A. T. A. (2005) Surface plasmon fluorescence measurements of human chorionic gonadotrophin: Role of antibody orientation in obtaining enhanced sensitivity and limit of detection. *Anal. Chem. 77*, 2426 – 2431.

Viitala, T., Vikholm, I., and Peltonen, J. (2000) Protein immobilization to a partially cross-linked organic monolayer. *Langmuir 16*, 4953 – 4961.

Wacker, R., Schroder, H., and Niemeyer, C. M. (2004) Performance of antibody microarrays fabricated by either DNA-directed immobilization, direct spotting, or streptavaidin-biotin attachment: a comparative study. *Anal. Biochem. 330*, 281 – 287.

Wang, H., Liu, Y., Yang, Y., Deng, T., Shen, G., and Yu, R. (2004) A protein A-based orientation-controlled immobilization strategy for antibodies using nanometer-sized gold particles and plasma-polymerized film. *Anal. Biochem. 324*, 219 – 226.

Weiss, G. A., and Sidhu, S. S. (2000) Design and evolution of artificial M13 coat proteins. *J. Mol. Biol. 300*, 213 – 219.

Westerlund, B., Van Die, I., Hoekstra, W., Virkola, R., and Korhonen, T. K. (1993) P fimbriae of uropathogenic Escherichia coli as multifunctional adherence organelles. *Zentrabl. Bakteriol. 278*, 229 – 237.

Westerlund-Wikstrom, B. and Korhonen, T. K. (2005) Molecular structure of adhesin domains in *Escherichia coli* fimbriae. *Int. J. Med. Microbiol. 295*, 479 – 486.

Wiersma, E. J., Collins, C., Fazel, S., and Shulman, M. J. (1998) Structural and functional analysis of J chain-deficient IgM. *J. Immunol. 160*, 5979 – 5989.

Wingren, C., and Borrebaeck, C. A. K. (2004) High-throughput proteomics using antibody microarrays. *Expert. Rev. Proteomics 1*, 355 – 364.

Wingren, C., and Borrebaeck, C. A. K. (2006) Antibody microarrays: current status and key technological advances. *Omics 10*, 411 – 427.

Wingren, C., and Borrebaeck, C. A. K. (2007) Antibody-based mircorarrays: From focused assays to proteome-scale analysis. *Chem. Mat. Sci. Bioarrays Pt. III,* 175 – 189.

Wingren, C., Steinhauer, C., Ingvarsson, J., Persson, E., Larsson, K., and Borrebaeck, C. A. K. (2005) Microarrays based on activity-tagged single-chain Fv antibodies: Sensitive detection of analyte in complex proteomes. *Proteomics 5*, 1281 – 1291.

Winzer, K., Falconer, C., Garber, N. C., Diggle, S. P., Camara, M., and Williams, P. (2000) The *Pseudomonas aeruginosa* lectins PA-IL and PA-IIL are controlled by quorum sensing and by RpoS. *J. Bacteriol. 182*, 6401 – 6411.

Wu, A. M., Gong, Y.-P., Li, C.-C., and Gilboa-Garber, N. (2010) Duality of the carbohydrate-recognition system of *Pseudomonas aeruginosa*-II lectin (PA-IIL). *FEBS Lett. 584*, 2371 – 2375.

Wu, S.-C., and Wong, S.-L. (2005) Engineering soluble monomeric streptavidin with reversible binding capability. *J. Biol. Chem. 280*, 23225 – 23231.

Xie, H., Guo, X.-M., and Chen, H. (2009) Making the most of fusion tags technology in structural characterization of membrane proteins. *Mol. Biotech. 42*, 135 – 145.

Yabe, R., Suzuki, R., Kuno, A., Fujimoto, Z., Jigami, Y., and Hirabayashi, J. (2007) Tailoring a novel silaic acid-binding lectin from a ricin-B chain-like galactose-binding protein by natural evolution-mimicry. *J. Biochem. 141*, 389 – 399.

Yamamoto, M., Kominato, Y., and Yamamoto. (1999) Phage display DNA cloning of protein with carbohydrate activity. *Biochem. Biophys. Res. Comm. 255*, 194 – 199.

Yeo, W.-S., Min, D.-H., Hseih, R. W., Greene, G. L., and Mrksich, M. (2005) Label-free detection of protein-protein interactions on biochips. *Angew. Chem. Intl. Ed. Engl. 44*, 5480 – 5483.

Zatta, P. F. (1996) A new bioluminescent assay for studies of protein G and protein A binding to IgG and IgM. *J. Biochem. Biophys. Meth. 32*, 7 – 13.

Zhang, Y., Li, Q., Rodriguez, L. G., and Gildersleeve, J. C. (2010) An Array-Based Method To Identify Multivalent Inhibitors. *J. Amer. Chem. Soc. 132*, 9653 – 9662.

Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bildingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R. A., Gerstein, M., and Snyder, M. (2001) Global analysis of protein activities using proteome chips. *Science 293*, 2101 – 2105.

Zinger-Yosovich, K., Sudakevitz, D., Imberty, A., Garber, N. C., and Gilboa-Garber, N. (2006) Production and properties of the native Chromobacterium violaceum fucose-binding lectin

(CV-IIL) compared to homologous lectins of Pseudomonas aeruginosa (PA-IIL) and Ralstonia solanacearum (RS-IIL). *Microbiology 152*, 457 – 463.

Zou, J., Glinksky, V. V., Landon, L. A., Matthews, L. and Deutscher, S. L. (2005) Peptides specific to the galectin-3 carbohydrate recognition domain inhibit metastasis-associated cancer cell adhesion. *Carcinogenesis 26*, 309 – 318.

Vita

Daniel Champlin Propheter was born in Santa Rosa, CA. Upon graduation from Jesuit High School, Carmichael, CA in May 2002, he attended Regis University in Denver, CO, and graduated in May 2006 with a Bachelor of Science. He then enrolled into the graduate program in the Department of Chemistry and Biochemistry at the University of Texas at Austin in August 2006, shortly after which he joined Dr. Lara Mahal's research group.

E-mail: proph388@gmail.com

This dissertation was typed by the author.